

**Oksana Petrova**  
**Of Pearls and Pigs**

**A Conceptual-Semantic Tiernet Approach to  
Formal Representation of Structure and Variation of  
Phraseological Units**





**Oksana Petrova**

Born 1977 in Saint Petersburg, Russia

M.A. (corr), Saint Petersburg state university, Russia, 2001

Doctoral student, Åbo Akademi University, Finland

Photo: Johan Paul

Cover: Tove Ahlbäck

Picture from Pieter Bruegel the Elder's *Nederlandish Proverbs*

Åbo Akademi University Press

Biskopsgatan 13, FI-20500 ÅBO, Finland

Tel. int. +358-20-786-1468

E-mail: [forlaget@abo.fi](mailto:forlaget@abo.fi)

<http://www.abo.fi/stiftelsen/forlag/>

Distribution: Oy Tibo-Trading Ab

P.O.Box 33, FI-21601 PARGAS, Finland

Tel. int. +358-2-454 9200

Fax int. +358-2-454 9220

E-mail: [tibo@tibo.net](mailto:tibo@tibo.net)

<http://www.tibo.net>

OF PEARLS AND PIGS





# Of Pearls and Pigs

A Conceptual-Semantic Tiernet Approach to Formal  
Representation of Structure and Variation of  
Phraseological Units

Oksana Petrova

ÅBO 2011

---

ÅBO AKADEMIS FÖRLAG – ÅBO AKADEMI UNIVERSITY PRESS

### **CIP Cataloguing in Publication**

**Petrova, Oksana.**

Of pearls and pigs : a conceptual-semantic  
Tiernet approach to formal representation  
of structure and variation of phraseological  
units / Oksana Petrova . – Åbo :

Åbo Akademi University Press, 2011.

Diss.: Åbo Akademi University.

ISBN 978-951-765-582-8

ISBN 978-951-765-582-8

ISBN 978-951-765-583-5 (digital)

Painosalama Oy

Åbo 2011

# Contents

<b>Acknowledgments</b>	<b>8</b>
<b>1. Introduction</b>	<b>9</b>
1.1 A brief terminological discussion	11
1.2 Research questions, objectives of the study and structure of the thesis	14
1.3 Computer-mediated discourse vs. traditional text corpora as a source of PU variation data	20
1.3.1 Traditional text corpora	21
1.3.2 Google Groups and Usenet	27
1.3.3 Obtaining PU variation data from Google Groups Usenet archive	30
1.3.4 Search results and categorization of data	34
<b>2. Previous approaches to some aspects of semantic description of PUs</b>	<b>40</b>
2.1 Semantic analysis of PUs in Russian phraseology	42
2.1.1 Phraseological meaning vs. lexical meaning	43
2.1.2 Broad meaning	45
2.1.3 The connotative component in the multi-component model of PU semantics	48
2.2 Semantic representation of PUs in Meaning-Text Model (MTM) Theory	57
2.3 Semantic representation of PUs in cognitive-linguistic models	69
2.3.1 Cognitive Grammar	69
2.3.2 Isomorphism and motivation as different dimensions of PUs' compositionality	77
2.3.3 Motivation by conceptual metaphors	82
2.3.4 Cognitive modelling of motivation in Conventional Figurative Language Theory (CFLT)	86
2.4 PUs and context	96
2.4.1 Psycholinguistic studies	96
2.4.2 Sufficient context, cohesion	96
2.4.3 Phraseological context	97
2.5 Negation and negative modality in PUs	101
2.6 Summary	107
<b>3. Formalized representation of PUs in the conceptual-semantic Tiernet model – theoretical and methodological aspects</b>	<b>110</b>
3.1 Conceptual Semantics and the Tiernet model	110
3.2 Organization of the conceptual structure and lexical linking rules	113
3.2.1 Linking types	113
3.2.2 Thematic tier	114
3.2.3 Thematic arguments (theta-arguments)	116
3.2.4 Action tier	117
3.2.5 Temporal tier	119
3.2.6 The semantic field tier (S-tier)	123
3.2.7 The modal tier	125
3.2.8 Principles of semantico-syntactic correspondence	127
3.2.9 Context and reference – the referential tier	128

3.3 Treatment of idioms and constructions within the framework of Conceptual Semantics	132
3.4 The network structure of PUs	135
3.4.1 Lexical constituents and their description	136
3.4.1.1 Tiers of phonological form	139
3.4.1.2 Tiers of morphophonology and morphosyntax	139
3.4.1.3 Tiers of the noun's Lexical Conceptual Structure (LCS)	141
3.4.2 Linking values and the default form of a PU	148
3.4.3 PU's conceptual structure and its linking to other levels of representation	157
3.4.3.1 The problem of literal vs. idiomatic meaning	157
3.4.3.2 Parallel conceptual structures in ambiguous PUs	159
3.4.3.3 The problem of analyzability	162
3.4.3.4 Reference in PUs	164
3.4.3.5 Unspecified thematic arguments and context dependency in the Finnish PU <i>HEITTÄÄ HELMIÄ SIOILLE</i> 'throw pearls to pigs'	174
3.4.3.6 Some remarks on the semantic field tier (S-tier) and the modal tier in PUs	180
3.5 Summary	187
<b>4. Critical overview of previous approaches to PU variation</b>	<b>189</b>
4.1 Problems in some previous syntactic and semantic accounts of PU frozenness and variability	191
4.2 Categorical models of variation – oppositions and classifications	196
4.2.1 Variation vs. transformations	197
4.2.2 Some major categories of variation and their definitional criteria	199
4.2.2.1 Formal vs. semantic variation	200
4.2.2.2 Usual vs. occasional variation	201
4.2.2.3 Systematic variation vs. wordplay	203
4.3 Phraseological blends	212
4.4 The notion of the base form and approaches to its definition	214
4.4.1 The lexicographic approach	215
4.4.2 The quantitative approach	218
4.4.3 The qualitative approach	221
4.4.4 The diachronic approach	222
4.4.5 Summary	224
4.5 The construction family and inheritance network approach to variation	224
4.6 Summary	245
<b>5. The Tiernet approach to the analysis of variation in PUs</b>	<b>247</b>
5.1 Principles of the Tiernet-based model of the construction family network	247
5.1.1 The notion of construction in the present approach	247
5.1.2 Constructions as networks	248
5.1.3 Variation as a construction family network	252
5.2 A PU's textual source, structure and discourse context as an underlying cause of variation	257
5.2.1 Association by virtue of lexical co-occurrence	259
5.2.2 Reference to discourse context	263
5.2.3 Textual dependence – relation between a PU and its source	267

5.3 Variation and the PU's network structure	273
5.3.1 Morphosyntactic patterns in the <i>HELMÄ SIOILLE</i> construction family	273
5.3.2 The effect of the modal tier on a PU's negative polarity	290
5.3.2.1 Morphosyntactic means of representing the PrNess→Neg features in discourse	291
5.3.2.2 Frequencies of negation	296
5.3.3 Default and non-default linking between a PU's syntactic constituents, phonology and LCS (lexical variation)	299
5.3.3.1 <i>Hopealusikka</i> 'silver spoon' or <i>satelliitti</i> 'satellite'? An example of weak default lexical linking	307
5.3.3.2 Non-default lexical linking licensed by a default constituent's phonological structure	314
5.3.3.3 Non-default lexical linking licensed by a default constituent's LCS	317
5.3.3.4 Non-default lexical linking licensed by the CS/PU	342
5.3.4 Interaction between constructions of different construction families	344
5.3.4.1 Host-donor interaction	344
5.3.4.2 Embedding and coordination	360
5.4 Discourse-related variation in PU	362
5.4.1 #SITUATION#/PU and its linking to the CS/PU	362
5.4.1.1 Economic relations	364
5.4.1.2 Interpersonal relations	366
5.4.1.3 Perception, experience	367
5.4.1.4 Structure and function	368
5.4.1.5 Verbal communication	369
5.4.2 Non-default lexical linking and the referential tier	370
5.4.2.1 Referentially licensed non-default nouns	370
5.4.2.2 Referentially licensed non-default verbs	372
5.4.3 Phraseological constituent isolation (PCI)	374
5.5 Summary	382
<b>6. Conclusion</b>	<b>384</b>
<b>Svensk sammanfattning</b>	<b>388</b>
<b>References</b>	<b>390</b>
<b>Abbreviations and symbols</b>	<b>412</b>

# Acknowledgments

*No man is an island.* It seems appropriate to start this book on phraseological units with a proverb which to the point describes the social nature of human life and activity in general and of scientific effort in particular. Just like social interaction is important for a human being, interaction within a scientific community is a vital part of any research.

Metaphorically speaking, a PhD project is a journey of a thousand miles and I am very pleased to thank people and institutions that helped me to carry on. I simply cannot thank enough my supervisor, Professor Urpo Nikanne at the Department of Finnish Language at Åbo Akademi University for being my guide and my scientific role model. I am also sincerely grateful to Pia Maria Ahlbäck who encouraged me to make the first step on this journey and to Päivikki Engblom who aided me through the final struggle. Many thanks to my companions on this journey – to my dear colleagues at the Department of Finnish Language for providing me with an outstandingly pleasant working environment and to fellow graduate students at Åbo Akademi University, Turku University and Finnish Graduate School in Language Studies, Langnet for creating a unique and inspiring atmosphere of scientific curiosity and endeavour throughout many courses, seminars and conferences which I had a privilege to attend.

I would like to thank the reviewers of this thesis, Professor Jan-Ola Östman and Professor Ray Jackendoff for their valuable critical remarks on the manuscript. Thanks to Anna Shepherd for her excellent work on the language check of the English text and to Robert Paul for checking the abstract in Swedish. The remaining flaws are all mine.

Working on this thesis was financially possible due to a doctoral student position at Åbo Akademi University Department of Finnish Language which I held from 2004-2009. I am also thankful to the Rector of Åbo Akademi University and to the Ella and Georg Ehrnrooth Foundation for funding parts of my research project. The Research Institute of the Åbo Akademi Foundation receives many thanks for having provided me with several travel grants, thus making it possible for me to attend some important conferences in my field of study.

Last, but not least, I would like to express my love and heartfelt gratitude to my mother Liudmila and to my husband Johan for their love, care and support.

Turku, 4.4.2011

Oksana Petrova

# 1. Introduction

There is an interesting claim made by Mel'čuk (1995: 169): "People do not speak in words, they speak in phrasemes." Without diminishing the role of regular syntax in natural languages we have to admit that phrasemes or phraseological units "constitute a large and central part of our knowledge of language and are valuable for the opportunities they afford for refining our understanding of how the abstract structures and principles of the human language faculty interact" (Tronenko 2003: 17).

In the present book, the focus of analysis are phenomena commonly referred to in linguistic literature as idioms or phraseological units (PUs) (see Section 1.1 of this chapter for a brief terminological discussion), i.e. conventionalized relatively stable multiword items of a given language exhibiting various kinds of irregularities in their structure. These items represent an interesting and challenging object of study. They are stored in speakers' memories, just like lexical items are, yet they are phrasal units. Although most of them obey the general syntactic rules of language, linking between syntax and semantic structure licensed by them is irregular. Their ability to undergo variation may be idiosyncratically constrained to a degree of complete fixedness, yet many of them permit variation, sometimes to an extent that blurs the distinction between a variant and a default form, so instead of one unit one is suddenly looking at a family of several interrelated constructions. Such a hugely variable, yet easily recognizable, form in which phraseological units are realized raises the difficult question of how these items are to be represented in language.

The purpose of this book is, on the one hand, by critically reviewing different previous approaches to describing phraseological units to make it very clear what difficulties their variation and variability raises for linguistic theory. On the other hand, in this thesis I attempt to find some new ways, both theoretical and methodological, in which these variable units can be approached. My goal has been to come up with a systematic and formalized description of phraseological units on the basis of their usage in written corpora. In order to achieve this, I have undertaken an extremely detailed usage-based study of the Finnish PU *HEITTÄÄ HELMIÄ SIOILLE* or 'throw pearls to pigs' – an equivalent of the English PU *CAST PEARLS BEFORE SWINE*. Thus, the study can be seen both in the context of

phraseological research and as a case study in applied corpus linguistics, although, as will be pointed out in Section 1.3, the asynchronous computer-mediated discourse, which was used as a source of data, differs from traditional corpora in many ways.

The theoretical framework, which I apply in my formal analysis, is Conceptual Semantics – a linguistic theory developed by Jackendoff (1972, 1976, 1983, 1987a, 1987b, 1988, 1990, 1993, 1994, 1995, 1996a, 1996b, 1997, 2002a, 2002b, 2002c, 2007a, 2007b and 2007c), Nikanne (1990a, 1990b, 1995, 1997a, 1997b, 1997c, 1998, 2000, 2002, 2005a, 2005b, 2006, 2008a and 2008b), Pörn (2005) and Paulsen (2011). It is important to mention that the version of Conceptual Semantics adopted here is the Tiernet model, developed by Nikanne (2002), which differs from Jackendoff's notational system. The formalism of Conceptual Semantics has proven to be efficient in treating irregularities in Finnish syntax, e.g. constructions where the choice of elative or ablative case is not rule-based (e.g. Nikanne 2005a); however, it has never been used for the purposes of phraseological analysis and thus needs further adjustments and development.

Following the methodological guidelines for Conceptual Semantics formulated in Nikanne (2008a), which include a formal approach, analytical organization, simple formation of modules and the importance of linking, the book features a very extensive, extremely detailed and increasingly complex formalism. Although I can sympathize with the reader, who might find the notation cryptic, there is very little I can do in this respect apart from gradually introducing and explaining separate parts of formal description and combining them into a vast network structure only at a later stage. So, I ask the reader to be patient and allow me to explain the new model of analysis throughout the many chapters of this book. The notation in its present form is significant, on the one hand, due the nature of the described phenomenon – phraseological units license complex structures. On the other hand, it serves the research goal, which is to show how these units work in their entire complexity, as a totality of links between different levels of representation.

The rigorous and stringent application of the above-mentioned guidelines resulted in several novel theoretical and notational solutions. Among the more general ones, one could name an elaborate system of linking types and values, a network representation of the temporal tier, a network representation of morphological structure and incorporation of



the referential tier into the Tiernet model. The more specific ones, developed especially for the purpose of phraseological analysis, are parallel conceptual and referential structures in ambiguous phraseological units, rejection of the distinction between literal and idiomatic meaning, the notion of the default form of a phraseological unit, and the new network approach to representation of the construction family, constituted by the default form and non-default variants.

## **1.1 A brief terminological discussion**

In the present study a general term phraseological unit and its abbreviation PU will be used to denote analyzed items. However, it is by no means the only term that can be found in literature on this subject. A significant terminological divergence exists in different fields of research that try to describe roughly the same phenomena. Used terms include idiom, phraseologism, phraseological unit, phraseme, formulaic expression, fixed expression, set phrase, multiword unit, multiword expression and phrasal lexical item, just to name a few and none of them being universally accepted (Wray 2002, Cowie 1998, Moon 1998, Everaert & Kuiper 1996). The choice of term is to a great extent a matter of tradition. For instance, the term phraseme is commonly used in international research on phraseology as a hypernym for all types that belong to the field of phraseology: idioms, similes, proverbs, restricted collocations, grammatical PUs and phrasal verbs. The umbrella terms phraseologism and phraseological unit are still preferred in Russian phraseology (e.g. Avdeeva 2004, Solodub & Al'brecht 2003, Dobrydneva 2000), but may sound unnatural in English. In Anglo-American tradition idiom is often used as a generic term for different kinds of PUs (Makkai 1972, Sinclair 1991, Fraser 1970, Fernando 1996). For argumentation in favour of the term phraseme over other terms, as well as discussion on the inappropriateness of using idiom as an umbrella term, see Dobrovol'skij & Piirainen (2005: 30). Naciscione (2001: 5) also argues that idiom is an ambiguous and loose term. Idioms in their narrow sense (in the Russian tradition also called phraseologisms or phraseological units proper) are often regarded as the central and most important class of PUs and, therefore, the core of phraseology (Dobrovol'skij & Piirainen 2005, Teliâ 1996, Kunin 1996). The main

argument underlying this idea is that idioms are the most irregular category, i.e. features of idiomaticity (irregularity of semantico-syntactic linking) and stability (inability to undergo certain formal variations) are manifested in them with a higher degree than in other PUs.

Terminological confusion is accompanied by the divergence in definitions given by different authors. Definitions may vary depending on which subclasses are chosen by the author to be included in the class of PUs. For example, cf. Kunin's (1996) definition of phraseological units: "A phraseological unit is a stable combination of words with a fully or partially figurative meaning", which excludes units with non-figurative meaning, with the statement made by Dobrovol'skij & Piirainen (2005: 29-30) that the term phraseme "is suitable to cover all kinds of conventional multiword units, figurative as well as non-figurative ones". In Dobrovol'skij & Piirainen (2005: 31) phrasemes are defined as "conventional polylexical units of the lexicon showing various kinds of formal and semantic irregularities". They are also assumed to be relatively stable in form and meaning.

Another definition of phrasemes can be found in Mel'čuk<sup>1</sup> (1995). Firstly, he points out that since phrasemes are phrases, their definition should be based on the concept of phrase, and thereupon he distinguishes phrases of two different types: a huge, theoretically unlimited class of free phrases and a very large but limited class of set phrases, or phrasemes (Mel'čuk 1995: 173). For free phrases it is true that "the signified, the signifier, and the syntactics are constructed exclusively according to the general rules of the language; a free phrase is thus 100% compositional and replaceable by any other sufficiently synonymous phrase". In other words the signified and the signifier of a free phrase are both unrestrictedly and regularly constructed, where unrestrictedness refers to "selection rules of a language" and regularity to "combination rules of a language". On the contrary, a phraseme is a phrase whose signifier and signified cannot be constructed both regularly and unrestrictedly (Mel'čuk 1995: 175).

The notion of collocation will be briefly reviewed in Section 5.2.1. It will be pointed out that collocation is an ambiguous term. Taken in the Firthian sense of collocation, i.e. as a habitual co-occurrence of

---

<sup>1</sup> A more detailed overview of Mel'čuk's approach to phraseology seen through the prism of his Meaning-Text Theory will be presented in Chapter 2 Section 2.2

syntagmatic items, PUs and collocations are phenomena of the same kind. If syntagmatic relations are seen as a key to the notion of a fixed expression, collocations and PUs are indeed synonyms. However, there is an alternative view widespread in the theory of phraseology, according to which collocations are seen as a subclass of PUs. In this tradition, collocations are defined as a semantic relation between two or more regularly co-occurring words, in which one of the words appears in its 'basic' meaning, while the other exhibits some combinatorial restrictedness, e.g. the adjective *stark* in *stark naked*. In order to avoid terminological confusion, I prefer not to use the term collocation while referring to analyzed items throughout this study.

Yet another term in need of clarification at an early stage of the study is the notion of construction. Here I will primarily use this notion in the sense of Nikanne (2005a: 199), who defines constructions as linking devices that license irregular linking patterns. In this respect, treatment of constructions in Nikanne's version of Conceptual Semantics is akin to that of Goldberg (1995). However, as opposed to Construction Grammars, in Conceptual Semantics all linking is not assumed to be governed by constructions. Irregular constructions are kept apart from regular syntactico-semantic linking patterns (Nikanne 2005a: 198). In this book, the notion of construction and the notion of PU in its most general sense are assumed to overlap. The choice of the term phraseological unit (PU) over the term construction was made partially as a tribute to the theory of phraseology and partially in order to avoid any possible confusion of the present study with Construction Grammars. Thus, the term PU will be used to denote an analyzed item as a lexicalized linguistic structure in general without reference to any particular variant, while the term construction will refer to any specific linking pattern, which can be regarded as a token of the given PU. Variation in a PU is, therefore, seen as a combination of its linking patterns. These are referred to as a PU's construction family.

Finally, the semantics/pragmatics distinction adopted in this book is akin to the relevance-theoretic cognitive approach introduced by Carston (2002), i.e. it is seen as a distinction correlating closely with two types of cognitive process: decoding and inference. Such distinction, in fact, is close to the distinction between two types of cognitive performance. The semantic representation of a phrase or an utterance is an output of a series of mappings licensed by an autonomous linguistic

system; it is encoded in what is uttered and decoded by the system. The pragmatic representation, on the other hand, is an output of the pragmatic inferential process, which “integrates the linguistic contribution with other readily accessible information in order to reach a confirmed interpretive hypothesis concerning the speaker's informative intention” (Carston 2002: 11). It is constrained and guided by the communicative principle of relevance. The pragmatics module is by no means considered to be a “dustbin”, but since the description of semantic structure is the main object of my concern here, pragmatic representation remains largely outside the scope of this study. Seeking the point of contact between the language faculty and the pragmatics module in the formal description of PUs could be a topic for future research.

The present study is to a very small extent concerned with terminological matters. As I see it, a “proper” labelling of a phenomenon comes secondary to its thorough, usage-based, explicit formal analysis, which is my primary objective here. In the following Section (1.2) I will dwell on the objectives of this study in more detail.

## 1.2 Research questions, objectives of the study and structure of the thesis

The prevailing tendency in phraseological research has been to draw assumptions about PUs, their meaning, structure and behaviour by looking at them in general and comparing them to each other without attempting to take a closer look at any PU in particular and describe it thoroughly. I believe that as long as our knowledge of PUs is superficial any kinds of generalizations will not be free of inadequacies. A detailed in-depth analysis of one single PU can in fact tell us more about the phenomenon and become a ground for further generalizations. Keeping that in mind, I have decided to focus my usage-based formal analysis mostly on one specific Finnish PU – *X HEITTÄÄ HELMIÄ SIOILLE* |  $\langle NP_{SUBJ}[N\{X\}] > V\{heittää\} \text{ ‘throw’} NP_{OBJ}[NP_{PL}\{helmi\} \text{ ‘pearl’}] PP_{ALL}[NP_{PL}\{sika\} \text{ ‘pig’}]$  | lit. ‘X throws pearls to pigs’, id. ‘X causes a transfer of some entity, evaluated by the speaker as good, to a recipient, evaluated as inadequate in some way’. This should not sound limiting as during the course of my study I refer to many more PUs, though not with the same depth of analysis. While I concentrate on this single PU, for which I have collected

extensive corpus data, many other phenomena will inevitably be drawn in.

By calling this study usage-based, I assume that it follows the criteria mentioned below:

- One of the goals for the study is to account for language use, i.e. how present-day Finnish speakers actually use the PU in question.
- Formal descriptions and theoretical implications presented in this book are developed on the basis of data gathered from authentic written discourse of modern Finnish.

My intention and my greatest challenge has been to be true to the data while developing a system of analysis that would be compatible with all tokens of PU variation detected in the data. Although the linguistic structure of the PU remains the main focus of this study, other relevant phenomena, such as intertextuality, allusions, context and cohesion, are inevitably drawn in and will be discussed in Sections 5.2.2, 5.2.3 and 5.4. Phenomena that did not receive due attention in this study are by no means irrelevant, but the large amount of interesting evidence found in the data, on the one hand, and the extremely detailed nature of analysis, on the other, made it virtually impossible to cover systematically the data in its entirety. Firstly it concerns phenomena of the semantic-pragmatic interface, mostly due to the fact that it would require developing an additional theoretical and notational system that would be compatible with the general conceptual-semantic theory and its methodological guidelines (Nikanne 2008a). The Tiernet-based model developed in this book is open in the sense that, in addition to those described here, it allows also other levels of linguistic representation to be connected with already existing tiers. In order to do that, however, one needs strictly defined primitives of each newly added level, as well as principles of their combination and their mapping to other levels. All this was a task impossible to accomplish within the limits of one book.

The present study is to a large extent aimed at finding regularities in the behaviour of this PU and constraints on its variation. I attempt to achieve it by studying 588 tokens of its use in the context of asynchronous computer-mediated discourse, namely in postings on Usenet newsgroups, which nowadays are available via Google Groups<sup>2</sup>. Since

---

<sup>2</sup> For a more detailed overview of this data-source and methods of data extraction see Section 1.3 below.

there has been practically no serious investigation of the use of PUs in computer-mediated communication, the present study attempts to fill in this gap. Usenet texts were expected to provide better evidence of PU variation than traditional corpora due to the archive's extensive size and the creative language use characteristic of this type of discourse. The data on which the study is based will be presented more fully in Chapter 5, although various cases are alluded to throughout the book. However, it is necessary already at this point to state clearly the problems that the data presents. Constructions found in the data can exhibit:

1. Variation in verb, as in *SYÖTTÄÄ HELMIÄ SIOILLE* 'feed pearls to pigs'.
2. Variation in nouns, as in *HEITTÄÄ HELMIÄ PORSAILLE* 'throw pearls to piglets' or *AJATUS ON HEITETTY SIOILLE* 'a thought has been thrown to pigs'.
3. Absence of verb, as in *HELMIÄ SIOILLE* 'pearls to pigs'.
4. Variation in thematic roles, as in *TARJOTA SIKOJA HELMILLE* 'offer pigs to Helmi' or *SIAT HEITTÄVÄT HELMIÄ LANTALÄTÄKKÖÖN* 'pigs throw pearls into a puddle of dung'.
5. Variation in linear order, e.g. *SELLAISTA HELMEÄ EMME SIALLE HEITÄ* 'such<sub>PTV SG</sub> pearl<sub>PTV SG</sub> NEG<sub>PL1</sub> pig<sub>ALL SG</sub> throw<sub>PRES ACT NEG</sub>' 'we will not throw such a pearl to a pig'.
6. Variation in case and number realization, e.g. *HEITTÄÄ HELMEN SIOILLE* 'throw **pearl**<sub>GEN SG</sub> pig<sub>ALL PL</sub>'.
7. Variation in syntactic realization, e.g. *HELMIEN HEITTÄMINEN SIOILLE* 'throwing (of) pearls to pigs'.
8. Blending with other PUs, e.g. *HEITTÄÄ HOPEALUSIKOITA SIOILLE* 'throw **silver spoons** to pigs' as a result of interaction with another Finnish PU *X YMMÄRTÄÄ/TIETÄÄ Y:STÄ YHTÄ PALJON KUIN SIKÄ HOPEALUSIKASTA* | *NP<sub>SUBJ</sub>[N{X}] V<sub>cogn</sub>{ymmärtää 'understand'/tietää 'know'} PP<sub>ELA</sub>{Y} COMP{yhtä paljon kuin 'as much as'} NP<sub>SUBJ</sub>[N{sika 'pig'}] PP<sub>ELA</sub>[NP{hopealusikka 'silver spoon'}]* | lit. 'X understands Y /knows about Y as much as a pig about a silver spoon', id. 'X does not understand Y/know about Y at all'.
9. Variation in semantics.

In connection with semantic variation, a few words have to be said about my paraphrase of the principal PU in the thesis as 'X causes a transfer of some entity, evaluated by the speaker as good, to a recipient, evaluated as inadequate in some way'. For someone who is not so familiar with my

data, it might seem that a more appropriate paraphrase would be something like ‘X transfers some valuable entity to a recipient who cannot appreciate its value’, which is also a common dictionary definition of this PU. However, the transferred entity is not good or valuable per se. This evaluative property is subjective – it is ascribed to the entity by the speaker whose point of view is not necessarily shared by all participants in the discourse. The same goes for the recipient’s inadequacy – the speaker’s view on this matter very seldom reflects the recipient’s opinion. As for the properties, I have deliberately chosen ‘good’ and ‘inadequate’, which might appear to be too general, instead of ‘valuable’ and ‘lacking appreciation’, which are unable to capture semantic variation in the data. For example, in (1) below the referent of *pearls* is a device (judging from a broader context a router) which possesses some attractive features, but home users and small companies do not make full use of these features. Do these features alone make this device ‘valuable’? Is it the ‘lack of appreciation’ that prevents an ordinary user at home from realizing the full potential of the device? My answer is – not necessarily. Thus, in order to provide a sufficiently unspecified description of this PU’s semantics, ‘good’ and ‘inadequate’ have been chosen. The former property is assumed to capture positive evaluation by the speaker, while the latter serves as some sort of generalization for different realizations that occur in the data for this PU (see Figure 64 in Section 3.4.3.5).

- (1) *Tavallisella kotikäyttäjällä tai pikkufirmassa se laite on lievästi ilmaistuna "helmiä sioille".*<sup>3,4</sup>

lit. ‘For an ordinary home user or in a small company this device is, to put it mildly, “**pearls to pigs**”’

Taken separately, the phenomena listed in 1–9 above can be accounted for within approaches in the literature. For instance, the classic base form – variant approach treats variation as deviation from the base form and attempts to classify variants alongside different lexical, morpho-syntactic and semantic-pragmatic parameters. The constructional instance – subpart approach, on the other hand, relates variants via taxonomic inheritance networks. However, in actual language data it is not uncommon for several different phenomena to be present within the

<sup>3</sup> <http://groups.google.com/group/sfnet.atk.sodat/msg/28a5370227f42881>, Jul 1 1999

<sup>4</sup> Further on in this book, all addresses for Google Groups (GG) messages will be shortened by omitting the <<http://groups.google.com/group/>> part.

same token. Neither classifications nor inheritance networks are capable of describing variants in all their complexity and diversity, simultaneously capturing their interrelations and differences. This is what I consider to be the real challenge for any kind of usage-based idiom variation analysis.

So the fundamental problem is how this PU is represented in the lexicon so that this variation can exist. This brings me to another aspect of PUs, which has not been paid due attention: the explicit formal description of their semantic structure and variation. There has been little if any attempt to formalize the findings of phraseological theory. Thus, although the results of the present usage-based study could also have lexicographic implementations, its primary goals rather lie within the scope of theoretical methodology; it is an attempt to adapt the theoretical model and formal descriptive tools of Conceptual Semantics for the purpose of integrated analysis of the PU's structure and variation. The present dissertation is the first attempt at a systematic and usage-based study of PUs and their variation in the framework of Conceptual Semantics. Therefore the general theoretical and methodological tasks are:

- To develop the theoretical model of Conceptual Semantics so that it can be applied to the comprehensive analysis of PU structure and could also be used describe PU variation. At the same time, this theoretical model has to be integrated into and be compatible with the general conceptual-semantic theory.
- To develop descriptive tools of Conceptual Semantics into a valid model for efficient and explicit formal description of PUs and other linguistic structures. An efficient model of semantic description should account for all kinds of correspondences between linguistic representations. The main desideratum is a formal model that could adequately deal with both regular and irregular linking, including cohesive linking to the discourse context.

The specific practical task – to give a thorough description of the selected Finnish PU – has the following subtasks:

- To analyze variation in the data.
- On the basis of this analysis to develop an explicit formal description of relevant aspects of the PU's structure, by linking together different parts of phonological, morphological, syntactic and conceptual representations.



- To apply this description to the analysis of underlying mechanisms in the PU's behaviour (regularities and restrictions in variation).
- To investigate contextual aspects of PU variation in relation to its structure, e.g. to apply the formal description of the PU's structure to the analysis of textual cohesion (such formalization of text description has never been attempted before).

The question of whether the PU's variants are all stored in speakers' memories or whether they are the product of some online capacity of accommodation will not be addressed in this study. It is most likely the case that PUs are stored in a form that permits variation, and the most frequent variants are probably stored as well, while infrequent ones are made up on the spot and integrated by hearers in terms of some less specific stored form. However, I will not make any commitments on this matter. Formal descriptions of the PU's structure, which are presented in the course of this book, are not aimed at explaining the online process of generating variants by speakers, or the process of listeners arriving at this structure. An account of how the structure is generated and derived from an underspecified lexicalized form would ultimately be a description of a process, while my task here is to describe a system which sets boundaries for this process. One of the major restrictive mechanisms introduced in this book is the relative strength of connections that make up the linguistic structure of PUs. At the same time, the formalism I develop here is able to capture the fact that any online interpretation of variants is secured by cohesive relations that exist between their structures on the one hand, and the discourse context in which they occur on the other. An empirical investigation of how such online interpretation actually happens in real time would demand methods different from those I use in this book, and is thus out of the scope of the present study.

This book is organized as follows. In Chapter 2 I will review previous approaches to some aspects of semantic description of PUs in different theoretical frameworks: Russian phraseology (Section 2.1), the Meaning-Text Model (Section 2.2) and cognitive-linguistic models (Section 2.3), such as Cognitive Grammar (Section 2.3.1), Conceptual Metaphor Theory (2.3.3) and Conventional Figurative Language Theory (Section 2.3.4). Also different approaches to aspects of PUs such as context (Section 2.4) and negation in their structure (Section 2.5) are examined. In Chapter 3 I will present an alternative approach to the description of PU structure in the framework of the conceptual-semantic Tiernet model. In Chapter 4 I will

survey different approaches to variation in PUs including different categorization models, the notion of the base form and the construction family and inheritance network approach in the sense of Goldberg (1995, 2006) and Croft (2001). Finally, Chapter 5 presents the Tiernet approach to the usage-based analysis of PU variation. The empirical data which has served as a basis for this analysis and methods for its collection are described in the following Section (1.3).

### **1.3 Computer-mediated discourse vs. traditional text corpora as a source of PU variation data**

The majority of previous research into idioms and phraseology is based upon dictionary entries, self-constructed examples, fictional literature or periodicals. Studies of the variation potential of PUs have been marred by a lack of authentic data or detailed examination of data. Nowadays scholars working in this field have gradually become more aware of the fact that PUs can only be properly described and understood if they are considered together with the contexts in which they occur, and this should involve corpus evidence (Moon 1998: 1). Such evidence questions many earlier statements about the syntactic behaviour of idioms. It is clear that variation is to be investigated more fully with corpora much larger than traditional ones (Moon 1998: 105). Findings from such corpus-based analysis might be useful for establishing new theoretical approaches to phraseology and linguistics in general. The role of context and cohesion in recognition and interpretation of altered PUs is extremely important. One could argue that, when too many constituents of a PU are altered or omitted, allusion to the PU is lost, and that examining a corpus cannot entirely reveal such cases since it requires the investigator to make up anomalous and ungrammatical examples. However, as will be demonstrated in Section 5.4.3, even a single isolated constituent can in principle be enough to establish the allusion, given that its context provides sufficient cohesive links to the PU's structure. Made-up examples can constrict the inquiry as any artificial context, which an investigator might construct in order to test the predictions of the analysis, will be inevitably biased.

In this section, which mainly follows discussion in Petrova (2010), I will approach some problematic aspects related to the extraction of PU variants from two different data sources, the first one being the

traditional text corpus represented by the Language Bank of Finland (Kielipankki) and the second one a type of asynchronous computer-mediated communication medium represented by Usenet discussion groups (Google Groups). By PU variation here I mean all kinds of discrepancies that can exist between the default morphosyntactic and phonological form of the target unit and the actual tokens of its occurrence in the data source. Some of this variation can be regarded as more or less regular and unrestricted, e.g. it is normally the case that verbs in predicate PUs inflect, although there may be some restrictions here as well. By relying on intuition or a smaller result set derived from a corpus, some types of variation (modification, transformation, substitution) can be perceived as more occasional, ad hoc manipulations with the default form, and others as non-available for a particular unit. However, examination of a larger result set can often lead to the discovery of unexpected variation phenomena and even shake our preconceived notions of variation and the default form. The natural way to obtain a larger result set is to look for a larger corpus, although there is no unequivocal answer to the question of how large a corpus should be in order to be considered “big enough”. The most obvious reason for this is the different frequencies of occurrence of different PUs within the same corpus, which can be big enough for a more frequent PU and not big enough for a less frequent one. For example, the Finnish PU *X KANTAA KORTENSA KEKOON* | *NP<sub>SUBJ</sub>[N{X}] V{kantaa ‘carry’} NP<sub>OBJ</sub>[N<sub>SG3</sub>{korsi ‘straw’}] PP<sub>ILL</sub>[N{keko ‘stack’}]* | lit. ‘X carries his/her straw to the stack’, id. ‘X does his/her bit’ with 276 hits returned by the query [bf='korsi'][[{0,2}[bf='keko']], while the entire result set for the Finnish PU *X HEITTÄÄ HELMIÄ SIOILLE* | *NP<sub>SUBJ</sub>[N{X}] V{heittää ‘throw’} NP<sub>OBJ</sub>[N<sub>PL</sub>{helmi ‘pearl’}] PP<sub>ALL</sub>[NP<sub>PL</sub>{sika ‘pig’}]* | lit. ‘X throws pearls to pigs’, id. ‘X causes a transfer of some entity, evaluated by the speaker as good, to a recipient, evaluated as inadequate in some way’ contains only 16 tokens. Thus, Usenet texts are expected to provide more substantial evidence of variation for less frequent PUs than traditional corpora do due to the archive’s extensive size and the creative language use characteristic of this type of discourse.

### 1.3.1 Traditional text corpora

By a traditional text corpus I mean a large and structured set of texts which are electronically stored, processed and often completed with

linguistic annotation (e.g. British National corpus, Russian National corpus, the Language Bank of Finland etc.). One of the main arguments in favour of using traditional corpora is that they “are carefully compiled in order to be used as a representative sample of language” (Hoffmann 2007) and therefore can be used to support generalizations concerning language use. Another argument is that they do not change very rapidly and the number of words can be determined at any time. This allows reliably replicating search results and counting normalized frequencies.

Rosamund Moon, who is one of the first linguists systematically to use corpus analysis in the study of PU variation, claims that “effective and robust descriptions of any kind of lexical item must be based on evidence, not intuition” (Moon 1998: 44). She points out that studies on the variation potential of idioms “are marred by a lack of authentic data or detailed examination of data” (ibid. 105). According to Moon, corpora would provide us with evidence of a suitable type. However, Moon (ibid.) herself admits that her corpus (an 18 million-word Oxford Hector Pilot corpus) is too small to give conclusive information about certain variations, and that variations have to be investigated more fully with much larger corpora. Given that my objective is a usage-based study of idiom variation in Finnish, the best available traditional corpus would be the Language Bank of Finland (Kielipankki) maintained at the Finnish IT Center for Science (CSC). It is Finland’s largest electronic corpus with approximately 130 million running words of Finnish texts, mostly periodicals from 1990-2000. Kielipankki is definitely larger than the OHPC used by Moon (1998). But does it provide enough examples of PU variation?

According to Moon (1998: 51), finding idiom variations is the hardest part of corpus-based investigations. Moon (ibid. 49) also emphasizes that the success of corpus investigation is entirely based on the effectiveness of the corpus tools. Traditional corpora offer a possibility of using powerful and linguistically oriented search syntax, and Kielipankki with its advanced search syntax is not an exception. However, even the most delicate and flexible corpus tools do not resolve all problems. One of these problems is that investigator bias can hardly be avoided: searches for PU variants are doomed to be deterministic and only report what has been sought, not what should or could have been looked for. Intuition is necessary, otherwise variations would not be found at all (ibid.). On the other hand, the construction of search queries preferably should not be

affected by preconceptions about non-variability of the investigated item, as this can possibly result in leaving some unexpected tokens of variation outside the search results. A similar point of view is expressed by Herold (2007: 61), who argues that one of the major principles for creating corpus queries is expecting all possible modifications. Thus, my assumptions about PU variation include at least the following points<sup>5</sup>:

1. Lexical constituents of a PU can appear in other syntactic constructions than the construction of the default form: e.g., the verbless construction *HELMIÄ SIOILLE* | *NP<sub>PTV</sub> PL[N{helmi 'pearl'} PP<sub>ALL</sub>[NP<sub>PL</sub>[N{sika 'pig'}]]]* | lit. 'pearls to pigs' instead of the transitive verbal construction *X HEITTÄÄ HELMIÄ SIOILLE* | *NP<sub>SUBJ</sub>[N{X}] V{heittää 'throw'} NP<sub>OBJ</sub>[N<sub>PL</sub>{helmi 'pearl'}] PP<sub>ALL</sub>[NP<sub>PL</sub>{sika 'pig'}]* | lit. 'X throws pearls to pigs'.
2. The linear order in which lexical components of a PU appear can differ from the word order of the default form: e.g., *X HEITTÄÄ SIOILLE HELMIÄ* | *NP<sub>SUBJ</sub>[N{X}] V{heittää 'throw'} PP<sub>ALL</sub>[NP<sub>PL</sub>{sika 'pig'}] NP<sub>OBJ</sub>[N<sub>PL</sub>{helmi 'pearl'}]* instead of *X HEITTÄÄ HELMIÄ SIOILLE* | *NP<sub>SUBJ</sub>[N{X}] V{heittää 'throw'} NP<sub>OBJ</sub>[N<sub>PL</sub>{helmi 'pearl'}] PP<sub>ALL</sub>[NP<sub>PL</sub>{sika 'pig'}]*.
3. The sequence of the default constituents can be interrupted e.g. by a modifier: e.g., *X HEITTÄÄ HELMIÄ SAASTAISILLE SIOILLE* | *NP<sub>SUBJ</sub>[N{X}] V{heittää 'throw'} NP<sub>OBJ</sub>[N<sub>PL</sub>{helmi 'pearl'}] PP<sub>ALL</sub>[NP<sub>PL</sub>[AP{saastainen 'filthy'}] N{sika 'pig'}]* | lit. 'X throws pearls to filthy pigs'.
4. Any default lexical component of a PU can appear in a morphological form that differs from that of the default form: e.g., *X HEITTÄÄ HELMIÄ SIALLE* | *NP<sub>SUBJ</sub>[N{X}] V{heittää 'throw'} NP<sub>OBJ</sub>[N<sub>PL</sub>{helmi 'pearl'}] PP<sub>ALL</sub>[NP<sub>SG</sub>{sika 'pig'}]* | lit. 'X throws pearls to a pig'.
5. The phonological/orthographic form of any of the PU's lexical components can differ from that of the default form: e.g., *X HEITTÄÄ HELEMIÄ SIOILLE* | *NP<sub>SUBJ</sub>[N{X}] V{heittää 'throw'} NP<sub>OBJ</sub>[N<sub>PL</sub>{helemi 'pearl'}] PP<sub>ALL</sub>[NP<sub>PL</sub>{sika 'pig'}]* | lit. 'X throws pearls to pigs', where the second *e* in *helemi* is a dialectal schwa-vowel.
6. The default lexical components can be substituted with other lexical items: e.g. *X SYÖTTÄÄ HELMIÄ SIOILLE* | *NP<sub>SUBJ</sub>[N{X}] V{syöttää 'feed'}*

<sup>5</sup> Here and henceforward the Finnish PU *X HEITTÄÄ HELMIÄ SIOILLE* | *NP<sub>SUBJ</sub>[N{X}] V{heittää 'throw'} NP<sub>OBJ</sub>[N<sub>PL</sub>{helmi 'pearl'}] PP<sub>ALL</sub>[NP<sub>PL</sub>{sika 'pig'}]* | lit. 'X throws pearls to pigs', id. 'X causes a transfer of some entity, evaluated by the speaker as good, to a recipient, evaluated as inadequate in some way' will serve as the primary example.

$NP_{OBJ}[NP_{PL}\{helmi\ 'pearl'\}] PP_{ALL}[NP_{PL}\{sika\ 'pig'\}]$  | lit. 'X feeds pearls to pigs'.

It is important to mention that the above assumptions are not to be considered as some special variation classes: they can both manifest themselves separately and freely combine within the same token, e.g. substituted components can appear in a non-default syntactic construction and also have non-default morphological forms.

When studying PU variation, the task of the corpus search would be locating possible PU variants with high accuracy. Accuracy is usually characterized by two aspects: precision (i.e. the query has to be composed in such a manner that it does not return too many irrelevant hits) and recall (i.e. the search does not miss too many relevant tokens). It is well known that maximizing recall typically leads to low precision, i.e. the fewer features a query specifies, the more relevant tokens are likely to be included, but at the same time a larger number of irrelevant hits will be returned, which in turn will demand manual analysis of a very large amount of data.

Herold (2007) remarks that lexical substitution is one of the major modifications that need to be taken into consideration during the query design. According to Herold (2007: 61), "we need to assume lexical substitution to be possible for every constituent". The same possibility for each lexical constituent to undergo substitution is expressed in Assumption 5 above. Thus, tokens where all three constituents of the Finnish PU  $X\ HEITTÄÄ\ HELMIÄ\ SIOILLE$  |  $NP_{SUBJ}[N\{X\}] V\{heittää\ 'throw'\} NP_{OBJ}[NP_{PL}\{helmi\ 'pearl'\}] PP_{ALL}[NP_{PL}\{sika\ 'pig'\}]$  | lit. 'X throws pearls to pigs', id. 'X causes a transfer of some entity, evaluated by the speaker as good, to a recipient, evaluated as inadequate in some way' are substituted are theoretically possible, but apparently they can be recognized as PU variants only in cases of substitution by lexical units bearing a close semantic relation (synonymy, hypernymy, hyponymy etc.) to the default ones (e.g. *nakata jalokiviä porsaille* 'toss precious stones to piglets'), or if the context contains the default constituents which cohere with the variant, e.g. (2) below:

(2) [...] *ennakkovaikutelmat tulevat koetun pohjalta, enkä ole vakuuttunut vielä japanilaisen sarjakuvan/animaation ihanuudesta, mutta jos voit heittää mielestäsi joitakin helmiä, niin nakkaappa tälle karjulle jokunen...*<sup>6</sup>

<sup>6</sup> GG: sfnet.harrastus.kulttuuri.sarjakuvat/msg/fe50541c78c3f5ed, May 11 1998

lit. '[...] preconceived impressions come on the basis of experience, and I am not yet convinced of the beauty of Japanese comics/animation, but if you think that you can **throw** some **pearls**, then **toss a few** to this **boar**...'

Since Kielipankki does not support queries based on semantic criteria, a query able to match variants with triple substitution has to be based on morphosyntactic criteria only:

```
[pos="Verb"][]{}{0,2}[pos="Noun"case="Part"number="PL"][]{}{0,2}[pos="Noun"
case="All"number="PL"]
```

In Kielipankki such a query returns 1299 hits. Possible relevant tokens can only be excerpted from this search result manually.

Tokens where any two of the three lexical constituents are substituted can be located by running a separate query for each constituent (the verb *heittää* 'throw' and the nouns *helmi* 'pearl' and *sika* 'pig'). The search query `[bf="heittää" pos="Verb"]`, which looks up tokens matching only the base form feature of the verb *heittää* 'to cast, throw' with no restrictions on the morphological form whatsoever, returns 17292 hits. The query `[bf="helmi" pos="Noun"]` returns 3187 hits, and the query `[bf="sika" pos="Noun"]` returns 2705 hits. Again, the only way to find relevant hits is to scroll through all the results and extract relevant hits manually, which is obviously a very labour-intensive and time-consuming task. A very common solution used by corpus linguists in order to decrease the amount of data is to look at a randomly selected subset. However, this solution is hardly applicable to the analysis of PU variation due to the low PU frequencies: e.g. Moon (1998: 60) observes that over 70% of fixed expressions and idioms in her database have frequencies of less than 1 per million tokens<sup>8</sup>. There is a high probability that randomly selected hits will not match any tokens of PU use at all, apart from tokens of its variation.

Since Moon (1998: 51) remarks that searches are most successful when the query consists of two lexical words fairly close together, several

---

<sup>7</sup> In Kielipankki's advanced search syntax query expressions, a search parameter is denoted by square brackets. A required feature in a search parameter is denoted by an equals sign "=". The feature's name is given to the left of the sign and the required value to the right of the sign within citation marks: `[key="value"]`. The empty search parameter matches any token whatsoever. The keys *bf* and *pos* are abbreviations of 'base form' and 'part of speech' respectively.

<sup>8</sup> Another problem is that idiom frequencies can be difficult to assess in the first place. Corpora are quantified in terms of individual words, but word-based frequency counts are not ideal for fixed expressions and idioms that are multi-word units (Moon 1998: 57).

queries, each consisting of two lexical components of the Finnish PU X HEITTÄÄ HELMIÄ SIOILLE |  $NP_{SUB}[N\{X\}] V\{heittää\}$  'throw'  $NP_{OBJ}[N_{PL}\{helmi\}$  'pearl']  $PP_{ALL}[N_{PL}\{sika\}$  'pig'] | lit. 'X throws pearls to pigs', id. 'X causes a transfer of some entity, evaluated by the speaker as good, to a recipient, evaluated as inadequate in some way', were made. The query  $[bf='helmi']\{0,5\}[bf='sika']$  finds all matches of the lemma *helmi* 'pearl' with the lemma *sika* 'pig' occurring within a window of between zero and five arbitrary tokens. This query returns 16 hits with 100% precision, i.e. all of the hits are relevant. The reversed order query  $[bf='sika']\{0,5\}[bf='helmi']$  matched 1 hit, which, however, was not a relevant one. The query  $[bf='heittää']\{0,5\}[bf='helmi']$  returned 3 hits, one of which was relevant, but the same token was already matched by the query  $[bf='helmi']\{0,5\}[bf='sika']$ . The query  $[bf='helmi']\{0,5\}[bf='heittää']$  returned 1 hit, which was not relevant. The query  $[bf='heittää']\{0,5\}[bf='sika']$  returned 3 hits, 1 of which was relevant and 0 new. The query  $[bf='sika']\{0,5\}[bf='heittää']$  returned no hits. Thus, all the above queries matched altogether 16 tokens of the PU in Kielipankki. These queries could find tokens of variation where constituents are used in other constructions than the base form, tokens with constituents in morphological forms other than the canonical forms and tokens where constituents are in a reversed order. They could also detect tokens where one of the three lexical constituents is substituted. However, they could not match tokens with two substituted constituents. In fact, the Kielipankki search results matched zero tokens with substituted noun constituents and 3 tokens with verbal substitution: 2 transitive (*tarjota* 'offer', *heitellä* 'fling') and one intransitive (*kadota* 'disappear').

The main advantages of traditional corpora are their representativeness, ability to count normalized frequencies and flexible linguistically oriented search tools. It has been pointed out that the latter two do not provide any substantial help in PU variation analysis. As for representativeness, one must be careful making generalizations on the basis of data obtained from Kielipankki, since its texts are to a large extent representative only of a single type of discourse: newspaper articles<sup>9</sup>. Although Kielipankki is larger than the OHPC used by Moon (1998), the scarce number of tokens which I could retrieve for the PU X HEITTÄÄ

---

<sup>9</sup> Nenonen (2007: 215) remarks that the absence of equal corpora of modern colloquial Finnish at present makes the World Wide Web the best source for this kind of language.



*HELMIÄ SIOILLE* | *NP<sub>SUBJ</sub>[N{X}] V{heittää ‘throw’} NP<sub>OBJ</sub>[N<sub>PL</sub>{helmi ‘pearl’}] PP<sub>ALL</sub>[NP<sub>PL</sub>{sika ‘pig’}]* | lit. ‘X throws pearls to pigs’, id. ‘X causes a transfer of some entity, evaluated by the speaker as good, to a recipient, evaluated as inadequate in some way’ is obviously not enough to identify generalities (patterns) of its use. Thus, based only on 16 hits it is virtually impossible to determine the default form of this PU. For example, while dictionaries would normally list the transitive verbal construction as the default form, 50% of the total of 16 tokens obtained from Kielipankki are represented by the verbless construction *HELMIÄ SIOILLE* | *NP<sub>PTV PL</sub>[N{helmi ‘pearl’}] PP<sub>ALL</sub>[NP<sub>PL</sub>[N{sika ‘pig’}]]* | lit. ‘pearls to pigs’. Finally, although corpus linguistics should be based on observation rather than introspection, finding tokens of idiom variation is inevitably a matter of serendipity (Moon 1998: 51).

### 1.3.2 Google Groups and Usenet

The Web as a corpus has its advantages: it is machine readable, free, easily accessible and, more importantly for the study of low-frequency phenomena, exceptionally large. For comparison, the query “*helmiä sioille*” *group:sfnet.\**, which searches for the exact string *helmiä sioille* ‘pearl<sub>PL PTV</sub> pig<sub>PL ALL</sub>’ in all *sfnet* groups, performed in Google Groups on 05 March 2009 returned 287 hits, i.e. about 18 times more than the above-mentioned more flexible searches in Kielipankki did. This particular search has 100% precision. However, its recall is low since it cannot retrieve any case and number variants or lexical substitutes of the noun components *helmi* ‘pearl’ and *sika* ‘pig’. Thus, the search indicates that the quantity of data for this PU that can be obtained via Google Groups is substantially larger than that of Kielipankki.

But if one considers the whole World Wide Web as a single corpus that can be accessed via commercial search engines such as Google, one faces a number of serious limitations (Hoffman 2007: 151). The first one concerns reproducibility: in contrast to traditional corpora, the World Wide Web is of indeterminable size and, moreover, is constantly growing, i.e. no normalized frequencies can be counted. Search results are very unstable and replicability of linguistic findings is virtually impossible in the World Wide Web. The second limitation has to do with search flexibility. Commercial search engines, like Google, Yahoo, AltaVista etc., do not allow search algorithms available in traditional corpus tools. It makes the retrieval of data for linguistic purposes a far more difficult and

time-consuming enterprise. Only specially designed linguistic pre-/post-processing search engines like Webcorp, KWiCFinder or Linguist's Search Engine are able to present examples of word usage from the Web in a form somewhat suitable for linguistic analysis.

Hoffmann (2007) presents a number of solutions which could make the Internet more suitable for linguistic investigations. One of them is to restrict the object of study to a clearly defined subsection of the World Wide Web. Another solution is to create a local copy of data by downloading relevant Web pages, post-process them and search with corpus tools. According to Hoffmann, using smaller and tailor-made Web-derived corpora allows to expand the range of available data without compromising on the application of standard corpus methodology. Hoffmann himself creates such a specialized Internet-derived corpus from a selection of Usenet newsgroup messages.

Usenet (USEr NETwork) is a global, decentralized computer network communications system. It was conceived in 1979 and by the 1990s it had developed into the largest system of discussion groups (often called newsgroups) on the Internet. It consists of thousands of discussion groups – hierarchically and thematically organized forums that allow people to share their thoughts and opinions on just about every imaginable subject and comment on the postings of others (Hoffmann 2007). Names of discussion groups indicate the topics that are discussed, e.g. the group *sfnet.keskustelu.foreigners* from the Finnish Usenet hierarchy *sfnet* is intended for foreigners living in Finland or visiting the country (*keskustelu* is the Finnish for 'discussion'). Figure 1 below shows an example of a Usenet message posted on *sfnet.keskustelu.kieli*, a newsgroup devoted to discussion on languages and their use. Here, the author replies to a question about the synonyms of the verb *inttää* 'argue, insist' in his idiolect. The question from the topic-starting message appears before the answer as a quotation, marked by an angle bracket at the beginning of each line of the quoted text.

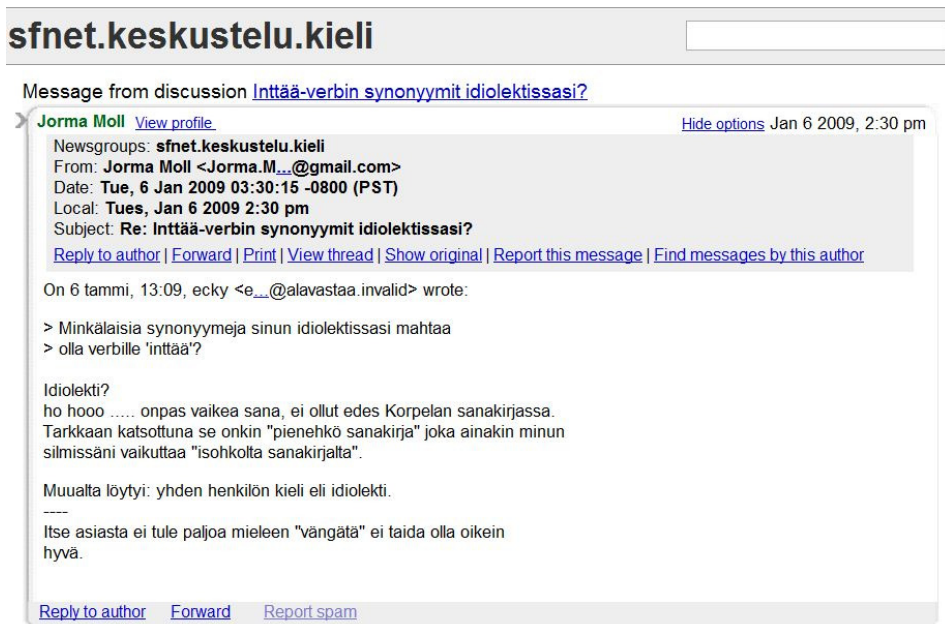


Figure 1 Screenshot of a Usenet message posted Jan 6 2009 on *sfnet.keskustelu.kieli/msg/00736490b46c901a*

Usenet newsgroups do not require participants to be online simultaneously, which puts them into the category of asynchronous computer-mediated communication (CMC). However, even though messages can be replied to with a considerable lapse of time, the nature of Usenet discussions is clearly interactive: participants often quote passages from previous posts as part of their replies, which greatly facilitates the establishment of topical coherence. Usenet discussions can thus be regarded as a hybrid form of communication, combining features of face-to-face talk with those of written texts (Hoffmann 2007). From the text-linguistic point of view, Usenet messages are a more homogenous data source than the World Wide Web. The interactive character of Usenet texts makes them an excellent data source for the study of discourse-pragmatic aspects of PU use and variation. The nature of the medium also stimulates its writers to use their language in an expressive, creative way, thus producing interesting tokens of occasional PU variation.

### 1.3.3 Obtaining PU variation data from Google Groups Usenet archive

The Usenet corpus-compiling procedure described by Hoffmann (2007) requires programming skills, UNIX system administration skills for setting up a news server, as well as access to a commercial newsfeed. The process of downloading the entire contents of all selected newsgroups onto the local hard disk can result in the transfer of enormous amounts of data and thus requires adequate hardware and network bandwidth. In the present study, I have chosen to obtain data via advanced searches in Google Groups<sup>10</sup>, which serve as the Web's most comprehensive archive of and interface to Usenet newsgroup postings dating back to 1981<sup>11</sup>. This method of data retrieval has its problems and limitations. However, taking into consideration the general problematical character of the application of corpus methodology for the study of PU variation discussed in the previous section, I will, in what follows, try to demonstrate that it can be justified, as long as I do not concern myself with normalized frequencies but concentrate primarily on the qualitative aspects of PU variation.

In previous corpus-based idiom variation studies (Moon 1998, Sköldberg 2004, Fellbaum 2007) data was gathered for a set of different expressions: Moon looks at a set of 6776 English fixed expressions and idioms, participants in the Wolfgang Paul-Preis Project whose results are presented in Fellbaum (2007) investigate some 1000 pre-selected German multi-word units, while Sköldberg restricts her set to 36 Swedish idioms. Due to the low frequencies of idioms in traditional corpora, the amount of tokens for the majority of idioms is rather low. For example, in Moon's data, 72% of fixed expressions and idioms have 0-17 tokens. Sköldberg, who deliberately chooses to look at idioms which occur in her 33 million word corpus with frequencies of more than one token per million words (i.e. not less than 33 tokens for a single idiom), reports that 32 in her set of 36 idioms are represented by less than 100 tokens, while the most frequent idiom is represented by 177 tokens. On the other hand, the above-mentioned test query *"helmiä sioille" group:sfnet.\** which returned 287 hits in Google Groups indicates that by running recall-maximizing

---

<sup>10</sup> <http://groups.google.com>

<sup>11</sup> Initially Usenet discussions were archived by DejaNews. Google acquired the archive in 2001.

queries the total number of tokens that could be obtained for the Finnish PU  $X$  HEITTÄÄ HELMIÄ SIOILLE |  $NP_{SUBJ}[N\{X\}]$   $V\{heittää\}$  'throw'  $NP_{OBJ}[N_{PL}\{helmi\}$  'pearl']]  $PP_{ALL}[NP_{PL}\{sika\}$  'pig']] | lit. 'X throws pearls to pigs', id. 'X causes a transfer of some entity, evaluated by the speaker as good, to a recipient, evaluated as inadequate in some way' can amount to several hundred. Such quantity of data allows one to perform a different kind of PU variation analysis: a thorough study of variation patterns in a single PU.

Thus, my goal was to gather variation data for  $X$  HEITTÄÄ HELMIÄ SIOILLE |  $NP_{SUBJ}[N\{X\}]$   $V\{heittää\}$  'throw'  $NP_{OBJ}[N_{PL}\{helmi\}$  'pearl']]  $PP_{ALL}[NP_{PL}\{sika\}$  'pig']] | lit. 'X throws pearls to pigs', id. 'X causes a transfer of some entity, evaluated by the speaker as good, to a recipient, evaluated as inadequate in some way'. Advanced search in Google Groups<sup>12</sup> allows restricting search queries according to several parameters: exact wording or phrase, language, site or domain, message date, group, subject and author. Since Google search engine does not offer a possibility of simultaneous searching for different word forms using wildcard truncation, separate search queries have to be carried out for all possible surface forms of the lexemes *helmi* 'pearl' and *sika* 'pig'. Searching for each and every lexical constituent separately in all possible forms is a very labour-intensive and time-consuming task. The situation is complicated by the fact that Finnish is a morphologically rich language; its nouns, verbs, and adjectives can theoretically have thousands of different inflected word forms (Karlsson 1983). Consequently, even if one restricts one's analysis to the case-number-possessive suffix paradigm, the noun components *helmi* 'pearl' and *sika* 'pig' will still have 312 different morphological forms altogether – 159 for the lexeme *helmi* 'pearl' and 153 for the lexeme *sika* 'pig'. Adding clitics (*-kin*, *-kAAn*, *-hAn*, *pA(s)*) would increase the number of query forms from 312 to 1872. Each dialectal or slang variant in combination with clitics would make this number even higher. In addition, as will be explained below, in order to increase the replicability of search results, for nine most frequent word forms separate searches were made for each year or a couple of years at a time, so, in point of fact, the actual number of searches was much higher than 312. The choice to exclude clitics, dialectal and slang variants, as well as cases of erroneous spelling, from the search queries was conditioned

---

<sup>12</sup> [http://groups.google.com/advanced\\_search](http://groups.google.com/advanced_search)

by practical reasons: most of these forms would return no relevant hits whatsoever, but the whole procedure of going through each and every one of them would be hideously time-consuming. Nevertheless, since the lexemes *helmi* ‘pearl’ and *sika* ‘pig’ were searched separately, such variants could be detected for each constituent, given that the other constituent appeared in its non-cliticized, standard-language version. Whenever detected, such cases were not excluded from the analysis.

The task at hand could have been made less complicated by using an inflection generator, on the one hand, and by creating custom-built software specifically fitting my research problem on the other. The latter kind of software would e.g. use Google SOAP Search API<sup>13</sup> to automatically search for and analyze different predefined inflection forms of a given word. Although this option was not used in the course of the current research, it is possible for designing such software to become a part of a future research project.

The development and implementation of custom-made software aimed at automatic extraction of PU variants is definitely a problem for future research, but it is not a problem that I aim to tackle here. To my knowledge, the problem of extracting idiom variants still remains largely unsolved in corpus linguistics. Herold (2007: 54) remarks: “Developing queries is essentially a manual task. We do not use techniques for automatic identification and extraction of target idioms or any other expressions.” He also points out that: “So far there is no sufficiently robust automatic process known to us that would permit corpus-driven extractions of idiomatic expressions” (Herold 2007: 56). There are indeed automatic processes that allow the extraction of statistically significant co-occurrences of certain lemmas, but I do not see how this can help anyone when it comes e.g. to the extraction of occasional lexical variants that cannot be predicted beforehand. As regards the application of part-of-speech annotation to the Usenet texts, it appears to be technically impossible due to the tremendous size of this archive and could only be achieved by downloading a very small part of it (as Hoffmann 2007 does).

Thus, when it comes to the detection of PU variants, the main object of my concern was to formulate queries that would meet the following principles: 1) all possible modifications have to be expected and 2) search accuracy has to be maximized so that all, or at least most, possible

---

<sup>13</sup> <http://code.google.com/apis/soapsearch/>

modifications can be retrieved. Queries corresponded to the following parameters:

- Search for exact wording or phrase (e.g. “*helmi*”, “*helmeni*”, “*helmesi*” etc.).
- Messages posted between 1 Jan 1990 and 31 Dec 2006.
- Messages from the group at this location: *sfnet.\**

Language restriction was unnecessary, since *sfnet.\** is by default a Finnish language hierarchy. In addition, a request was made to sort search results by date and display 100 results per page. The search results page presents the results in the following form:

**Karalahti pillittää taas**

*sfnet.keskustelu.vitsit* - 74 posts - 16 authors - Last post: Apr 30, 2006

... reader1.news.jippii.net *sfnet.keskustelu.vitsit*:178905 Jupehan se siellä tykittää taas. Tiesikö edes jupe tuota. Rietas naisenkuva on vain **helmi** sioille?

<http://groups.google.com/g/8fb7ff09/t/f448990d1507ea9a/d/bbe546cbddfcabaf>

Here, the first row is the title of the discussion thread (*Karalahti pillittää taas* ‘Karalahti cries again’). The second row contains the name of the discussion group (*sfnet.keskustelu.vitsit*), number of posts in the thread (74), number of authors (16) and date of the last post (30 Apr 2006). The following two rows constitute the snippet, which is Google’s algorithmic attempt to extract the part of the discussion thread most relevant to the search query. Normally it is an excerpt from the message containing the searched item in boldface type (**helmi** ‘pearl’ in the above example). In most cases the snippet is enough for determining whether the hit is relevant or not, and thus there is no need to open every thread returned by the search. The last row contains the thread’s URL address.

Making sense of Google search results is actually quite a difficult problem to tackle. For instance, the search for the exact form “*helmi*” on *sfnet.\** between 1 Jan 1990 and 31 Dec 2006 returns about 5,380 results. Firstly, this is by no means an exact number. For the sake of efficiency, Google estimates the number of results, and this estimate of the total number of results is rather unreliable. Secondly, Google would never display more than about 400-700 (presumably, randomly selected) search results. My solution to this problem was to run for nine most common morphological forms<sup>14</sup> separate searches for each year or a couple of years

---

<sup>14</sup> These kinds of queries were performed for the following word forms: “*helmi*”, “*helmet*”, “*helmiä*”, “*sika*”, “*siat*”, “*sian*”, “*sikaa*”, “*sikoja*”, “*sioille*”.

at a time and then sum up the number of results. Thirdly, search results only show the number of discussion threads where the searched item occurs, not the actual number of occurrences. Thus, considering that the same item can possibly occur several times within the same discussion thread, the actual number of occurrences can be much higher. Overall, having only Google search engine at hand, counting the exact number of occurrences for each word form seems to be a difficult, perhaps even impossible, task. On the other hand, it is unnecessary since my goal is limited to the retrieval of PU variants. The only object of my concern should be the reproducibility of search results. As has been pointed out earlier, search results obtained from the Web are unstable and therefore not reproducible due to the fact that Web content is constantly changing and web pages can disappear. In this sense, Google Groups as a corpus has one important advantage: old messages in the archive can still be retrieved and running advanced searches with restricted message dates could in theory be reproducible. However, unfortunately this is not exactly the case. Searches for one year at a time significantly improve reproducibility, but even they can return slightly different results on different occasions (usually about  $\pm 0.5$  hits) for reasons, which remain largely unknown since Google search algorithms are Google's trade secret.

### 1.3.4 Search results and categorization of data

Only 47 out of 159 word forms for the lemma *helmi* 'pearl' have returned any results and among these only 20 word forms have returned results containing altogether 496 relevant word form tokens. For the component *sika* 'pig', 42 out of 153 word forms have returned any results and only 14 word forms have returned results containing altogether 470 relevant word form tokens. By a relevant word form token I mean a single token of the searched word form occurring in a sentence which can, according to the semantic criteria, be regarded as a context of the PU *X HEITTÄÄ HELMIÄ SIOILLE* | *NP<sub>SUBJ</sub>[N{X}] V{heittää 'throw'} NP<sub>OBJ</sub>[NP<sub>PL</sub>{helmi 'pearl'}] PP<sub>ALL</sub>[NP<sub>PL</sub>{sika 'pig'}]* | lit. 'X throws pearls to pigs', id. 'X causes a transfer of some entity, evaluated by the speaker as good, to a recipient, evaluated as inadequate in some way' use. Query results were then exported into the Microsoft Access database application, where each record corresponds to a single token of PU use within an autonomous syntactic construction. For instance, the Example (3) below was recorded as two



separate entries (Table 1) – one containing the verb *poimia* ‘pick’ and the noun component *helmi* ‘pearl’ and the other including the verb *tarjota* ‘offer’ and the noun component *sika* ‘pig’:

- (3) *Mutta ajattelinkin että joku poimisi helmet ja tarjoaisi sioille, mina vaan röhnöttäisin sillä välin (kyljelläni) ja röhkisin.*<sup>15</sup>  
 lit. ‘But I thought that someone would **pick pearls** and **offer to pigs**, I would just loll about (on my side) and grunt’.

Table 1 Lexical constituents HELMI ‘pearl’ and SIKA ‘pig’ distributed between two autonomous syntactic constructions.

Date/Author/Group	Message body
Dec 16 1997/ Korhonen Tommi/ sfnet.keskustelu.seksi	<i>Mutta ajattelinkin että joku poimisi <b>helmet</b> [...] lit. ‘But I thought that someone would <b>pick pearls</b> [...]’</i> <i>[...] ja <b>tarjoaisi sioille</b>, mina vaan röhnöttäisin sillä välin (kyljelläni) ja röhkisin. lit. ‘ [...] and <b>offer to pigs</b>, I would just loll about (on my side) and grunt’.</i>

In the database different types of data relating to the organizational and formal aspects of individual tokens of PU use were recorded: type of construction, morphological form of each noun constituent (case, number and possessive suffix), negation and modality, lexical substitution, word order, modifiers, appellatives and evidentials. Organizational fields recorded message ID, date, year, author and the name of discussion group. The entire database contains 588 tokens of PU use occurring in 521 different newsgroup messages written by 343 different usernames<sup>16</sup> in 97 different sfnet.\* groups. The question of what should be considered a PU variant is not a trivial one. The default form of a PU is usually defined as a form which simultaneously meets several different criteria:

- Phonological criterion, i.e. the presence of certain lexical items in the same structure.
- Morphosyntactic criterion, i.e. a particular structure in which lexical items appear.
- Semantic criterion, i.e. a particular conceptual structure associated with the phonological and syntactic structures.

<sup>15</sup> GG: sfnet.keskustelu.seksi/msg/937ee650b2db8fa9, Dec 16 1997

<sup>16</sup> It is difficult to trace whether the same author is actually writing under several different usernames.

- Criterion of institutionalization, i.e. the string being recognized and accepted as a phraseological unit of the language, or in corpus terms the frequency of the string (Moon 1998).

The last criterion is neither necessary nor sufficient for a PU variant (although some variation classifications distinguish between usual vs. occasional variation). As for the first three (phonological, morphosyntactic and conceptual structures), the most difficult problem is to determine which combination of these would be necessary and sufficient for a variant to be considered as a member of the *HELMIÄ SIOILLE* construction family. Borderline cases are inevitable and one has to decide whether or not to include them into the database. My solution was to record them in the database as well, but by adding a technical field UNCLEAR enable their filtering from more clear-cut results. This was mainly done for the sake of the quantitative morphosyntactic analysis, which allowed distributing clear tokens between different patterns<sup>17</sup> without completely discarding interesting but less clear tokens of PU variation. Thus, 85 borderline tokens (marked as UNCLEAR in the database) have been recorded. These include quotations of the original biblical passage (also slightly inexact ones, like in Example (4) below, where *helmiä* ‘pearl<sub>PL PTV</sub>’ lacks a possessive suffix *-nne* ‘2PL’, which is present in the biblical source):

- (4) *Sillä tiedäthän, että "Älkää heittäkö helmiä sikojen eteen, ja sitä mikä on pyhää, koirille, etteivät ne kääntyisi ja repisi teitä".*<sup>18</sup>

lit. ‘For you know, that “**Do not cast pearls before swine** and what is holy to the dogs, lest they turn and tear you in pieces”’

Although such quotations meet both the phonological requirement and the semantic requirement, they cannot be considered as tokens of the PU in a strict sense. As Dobrovol’skij & Piirainen (2005: 231) remark, there are many text fragments that were initially used as citations before they gradually developed into conventional figurative units. In the above example both citation marks and inclusion of the *holy to the dogs* passage indicate that we are dealing with a biblical quotation, which should be considered as a source for this particular PU, rather than with the PU itself. Distinction between quotations and non-quotations had to be made primarily out of consideration for the results of morphosyntactic analysis,

<sup>17</sup> These patterns are presented in Chapter 5, Section 5.3.1.

<sup>18</sup> GG: sfnet.keskustelu.ihmissuhteet/msg/ee5db18c263f52d8, Nov 9 2003

the purpose of which was to investigate the recurring patterns of this PU in modern Finnish. If quotations were counted together with clear PU tokens, the pattern represented by quotations would have gained the status of a present-day construction, which is obviously not the case.

Another category of borderline tokens includes cases where both lexical components *helmi* ‘pearl’ and *sika* ‘pig’ are hosted<sup>19</sup> by another syntactic construction belonging to a different construction family<sup>20</sup>. For instance, in the Example (5) below the host is the Finnish PU *X ETSII/ SEULOO/ POIMII/ TONKII/ LÖYTÄÄ/... HELMEN/ HELMET/ HELMIÄ ROSKASTA/ PASKASTA/ ROMUSTA* | *NP<sub>SUBJ</sub>[N{X}] Vtr{etsiä ‘search’/seuloa ‘sieve’/poimia ‘pick’/tonkia ‘dig’/löytää ‘find’/...} NP<sub>OBJ</sub>{**helmi** ‘pearl’} PP<sub>SEPAR</sub>[NP{roska ‘garbage’/paska ‘shit’/romu ‘junk’}]* | lit. ‘X searches/sieves/picks/digs/finds/ etc. pearls from garbage/shit/junk’, id. ‘X searches/... some fine, valuable and rare entities from a bulk of useless stuff’, where NP constituent in PP<sub>SEPAR</sub> corresponds to *sika* ‘pig’, and in the Example (6) below the host is another Finnish PU *X YMMÄRTÄÄ/TIETÄÄ Y:STÄ YHTÄ PALJON KUIN SIKA HOPEALUSIKASTA* | *NP<sub>SUBJ</sub>[N{X}] Vcogn{ymmärtää ‘understand’/tietää ‘know’} PP<sub>ELA</sub>{Y} COMP{yhtä paljon kuin ‘as much as’} NP<sub>SUBJ</sub>[N{**sika** ‘pig’}] PP<sub>ELA</sub>[NP{hopealusikka ‘silver spoon’}]* | lit. ‘X understands Y /knows about Y as much as a pig about a silver spoon’, id. ‘X does not understand Y/know about Y at all’<sup>21</sup>, where NP<sub>ELA</sub> corresponds to *helmi* ‘pearl’. Boldfaced lexical items in both of the host constructions are shared with the *HELMIÄ SIOILLE* construction family and are preserved in the resulting blends together with original syntactic structure of the hosts. Since both *helmi* ‘pearl’ and *sika* ‘pig’ appear in the blend, one could assume that (5) and (6) below are tokens of *X HEITTÄÄ HELMIÄ SIOILLE* | *NP<sub>SUBJ</sub>[N{X}] V{heittää ‘throw’} NP<sub>OBJ</sub>[NP<sub>PL</sub>{**helmi** ‘pearl’}] PP<sub>ALL</sub>[NP<sub>PL</sub>{**sika** ‘pig’}]* | lit. ‘X throws pearls to pigs’, id. ‘X causes a transfer of some entity, evaluated by the speaker as good, to a recipient, evaluated as inadequate in some way’. Nevertheless, according to both the syntactic and the semantic criteria these are rather tokens of the above-mentioned host PUs inasmuch as they retain the original morphosyntactic and conceptual structure. It is

<sup>19</sup> The notions of host and donor constructions and the process of their interaction are described in Chapter 5, Section 5.3.4.1 of this book.

<sup>20</sup> The notion of construction family in the framework of Construction Grammar is discussed in Chapter 4, Section 4.5. A conceptual-semantic Tiernet approach to this notion is presented in Chapter 5, Section 5.1.3.

<sup>21</sup> In Chapter 5, Section 5.3.3.1 I take a closer look at this particular construction family.

worthwhile to mention that these tokens are relevant for our understanding of variational mechanisms and they will be analysed in more detail in Section 5.3.4.1 of this book. However, they were not counted among constructions of the *HELMIÄ SIOILLE* family.

- (5) *Joudut siis noukkimaan helmiä sikojen joukosta.*<sup>22</sup>

lit. 'So you have to **pick pearls among pigs**'

- (6) *Tiedätte epilepsiasta yhtä paljoa kuin sika helmistä!*<sup>23</sup>

lit. 'You know as much about epilepsy **as a pig about pearls.**'

On the other hand, tokens where one of the recurrent *HELMIÄ SIOILLE* constructions itself functions as a host for a lexical component from another construction family are not labelled as UNCLEAR and are therefore counted together with other similar constructions in the database. For example, in Example (7) below the component *helmi* 'pearl' is substituted by the lexical unit *hopealusikka* 'silver spoon' which is borrowed from the PU *X YMMÄRTÄÄ/ TIETÄÄ Y:STÄ YHTÄ PALJON KUIN SIKA HOPEALUSIKASTA* | *NP<sub>SUBJ</sub>[N{X}] Vcogn{ymmärtää 'understand'/ tietää 'know'} PP<sub>ELA</sub>{Y} COMP{yhtä paljon kuin 'as much as'} NP<sub>SUBJ</sub>[N{sika 'pig'}] PP<sub>ELA</sub>[NP{hopealusikka 'silver spoon'}]* | lit. 'X understands Y /knows about Y as much as a pig about a silver spoon', id. 'X does not understand Y/know about Y at all':

- (7) *Enpä taida enään herra Burmaniin soveltaa ironiaa, sehän on kuin hopealusikoita sioille.*<sup>24</sup>

lit. 'I am not likely to apply irony to Mr. Burman anymore, it is like **silver spoons to pigs.**'

Isolated lexical components i.e. components of a PU that do not occur together with other lexical components of the same PU within the syntactic structure of the same clause (Petrova 2007b) were also treated as borderline tokens, e.g. in (8) below:

- (8) *Eikä sioissakaan mitään vikaa ole. (Hengellisiä) helmiä kun on loputtomasti tarjolla. Toisaalta siat haluaisivat ennemmin ruokaa, koska siitä on heille enemmän hyötyä. Mutta helmet voivat ne tappaa joutuessaan henkireikään.*<sup>25</sup>

lit. 'There's nothing wrong with **pigs** either. For (spiritual) **pearls** are in endless supply. On the other hand, **pigs** would rather like food, because it is of more benefit to them. But **pearls** can kill them if they get into the air hole.'

<sup>22</sup> GG: sfnet.viestinta.nyyssit/msg/d6425463f4112183, Oct 31 1997

<sup>23</sup> GG: sfnet.keskustelu.varaventtiili/msg/8a6c3be4b14334d5, Sep 9 2005

<sup>24</sup> GG: sfnet.atk.sodat/msg/4d5ef2ebb5e7a858, May 22 1997

<sup>25</sup> GG: sfnet.keskustelu.uskonto/msg/4d2c63d56482e68f, Dec 24 1997

Isolated lexical items of the type presented in (8) above do not appear as constituents of any particular construction of the *HELMIÄ SIOILLE* construction family. However, they represent extremely valuable material for the study of phraseological cohesion and will be discussed in more detail in Section 5.4.3 of this book.

As for the cases where NPs *helmi* ‘pearl’ and *sika* ‘pig’ are distributed between the main clause and the relative clause as is shown in Table 2 below, although the main clauses containing modified NPs *helmi* ‘pearl’ and *sika* ‘pig’ (labelled as MODIFIED N1 or MODIFIED N2) have been recorded as separate tokens, they were excluded from the final calculation of construction patterns, where only tokens of the relative clause constructions have been counted according to the construction which they represent.

Table 2 *PU constituents distributed between the main clause and the relative clause*

Date/Author/Group	Message body	CONSTR
Jun 27 1998/ Patrick Uotinen/ sfnet.keskustelu.uskonto	<i>Gregorius, joka siis itse tallensi omat sanansa, jatkoi vielä letkautuksella <b>helmistä</b>, [...] lit. ‘Gregorius, who himself recorded his own words, continued with a quip about <b>pearls</b> [...]’</i>	MODIFIED N1
	<i>[...] <b>joita</b> hänen ei tarvitse <b>heitellä</b> saastaisen <b>sian eteen!</b><sup>26</sup> lit. ‘[...] <b>which</b> he does not need to <b>throw before a filthy pig!</b>’</i>	Vtr-N1 <sub>OBJ</sub> -N2 <sub>TERM</sub>

The remaining 480 tokens represent morphosyntactic patterns (constructions) discussed in Chapter 5, Section 5.3.1.

<sup>26</sup> GG: sfnet.keskustelu.uskonto/msg/c6ab0b13214142c4, Jun 27 1998

## 2. Previous approaches to some aspects of semantic description of PUs

Problems concerning different aspects of idiomaticity and idioms have been tackled by practically every linguistic theory from the structuralist framework (e.g. Hockett 1958; Healey 1968; Makkai 1972) to transformational generative frameworks (e.g. Katz & Postal 1963; Chafe 1968; Weinreich 1969; Fraser 1970; Wasow, Sag & Nunberg 1983; Nunberg, Sag & Wasow 1994; O'Grady 1998; Tronenko 2003), construction grammars (e.g. Fillmore, Kay & O'Connor 1988, Jackendoff 1997, Kay & Fillmore 1999; Penttilä 2006), corpus studies (e.g. Moon 1998; Stubbs 2001; Sköldberg 2004), psycholinguistic studies (e.g. Bobrow & Bell 1973; Swinney & Cutler 1979; Gibbs 1980; Cacciari & Tabossi 1988; Flores d'Arcais 1993; Peterson & Burges 1993; Cacciari & Glucksberg 1994), linguistic typology (Dobrovolskij 1988, 1992) and cognitive grammars (e.g. Langacker 1987; Dobrovolskij 1995; Langlotz 2006). A quite comprehensive overview of previous research on idioms and idiomaticity in the Anglo-American tradition can be found in e.g. Penttilä (2006). The majority of works that belong to this tradition have been reviewed previously on multiple occasions, and it seems unnecessary to present them once more in the scope of this book.

There is relatively little literature on idioms in Fennistics. A substantive part of research on phraseology in the Finnish language have been executed by Germanic linguists who have studied Finnish idioms as well (e.g. Hyvärinen 1996a, 1996c, Korhonen 2000), but mainly in contrast with German ones (e.g. Hyvärinen 1992a, 1992b, 1996b, 2004; Korhonen 1987, 1991, 1995, 1996, 1998). Hyvärinen (2007) presents a detailed review of works on idioms and phraseology in Finnish. During the 1950-60s Finnish proverbs were analyzed by M. Kuusi in several articles (e.g. 1952, 1963), where he int. al. discussed terminological issues. There are practically only a couple of monographs in this area, among them Järviö-Nieminen's (1959) account of Finnish wellerisms, A.-L. Kuusi's (1971) classification of phraseological units and a doctoral thesis by Nenonen (2002). Kuusi's work is actually not much of a scientific research but rather an attempt to classify Finnish idioms according to principles that prove to be quite inconsequent. Nenonen's research is based on empirical data collected from various corpora and from psycholinguistic

experiments. She deals with morphological, lexical and syntactic aspects of Finnish idioms and experimentally tests the syntactic restrictedness of VP idioms. By means of corpus analysis she detects some idiom-prone verbs and nouns in Finnish. The prototypical Finnish phrasal idiom, according to Nenonen, is a verb phrase that consists of a basic verb and an inflected noun. Morphological and psycholinguistic analysis of Finnish idioms can also be found in e.g. Nenonen (2001a, 2001b, 2007a, 2007b); Nenonen & Niemi (1999, 2004, 2010); Nenonen, Niemi & Laine (2000, 2002); and Niemi, Nenonen & Penttilä (1998). Some of the collected articles in Lähdemäki & Bertills (2000) concern Finnish idioms, although other languages are analyzed as well. The problems discussed are idiom definition (Häkkinen 2000), translation and equivalence (Ingo 2000), grammaticalization (Mikone 2000) and idioms in language use (Karlsson 2000).

A brief outline of phraseological research in Europe can be found in Dobrovol'skij & Piirainen (2005). The authors emphasize that the long tradition of phraseological research in the tradition of Bally (mostly written in Russian, German or French) is practically unknown to Anglo-Saxon linguists. Very little is known about the Russian phraseological theory, e.g. works of Vinogradov, Amosova, Akhmanova, Šanskij, Arhangelskij, Teliâ, V. Žukov, Kunin and Mel'čuk. As a native speaker of Russian I have an opportunity to somehow fill in this gap – Section 2.1 of the current chapter reviews several methods of semantic analysis of PUs in Russian phraseological theory.

This chapter as a whole is aimed at providing a critical overview of different models and approaches to semantic description of PUs within the framework of several linguistic theories. Since the literature on PUs is enormous, a complete survey would be an unrealistic option. Therefore, this overview has to be limited to only those aspects of semantic analysis that are relevant for the present study, i.e. the conceptual-semantic and Tiernet approach to the formal representation and variation of PUs presented in Chapter 3 and Chapter 5. The current chapter will start with Section 2.1, which looks at the methods of semantic analysis of PUs in Russian phraseological theory. It includes exploration of the postulated distinction between phraseological and lexical meaning (2.1.1), the notion of the broad meaning (2.1.2) and multi-component models of PU semantic structure, which emphasize the importance of connotation within this structure (2.1.3). Section 2.2 presents the framework of the

Meaning-Text theory, which is the first attempt to develop a system of formal description of collocations (in their phraseological sense). Section 2.3 evaluates cognitive-linguistic models of semantic representation of PUs: cognitive grammar (2.3.1), a prismatic model of isomorphism and motivation as different dimensions of PUs' compositionality (2.3.2), idiom motivation by conceptual metaphors (2.3.3) and cognitive modelling of motivation in Conventional Figurative Language Theory (2.3.4). Section 2.4 examines several approaches to the problem of PUs' relation to their context. Finally, Section 2.5 presents a brief overview of several studies of negation and negative modality in PUs.

## 2.1 Semantic analysis of PUs in Russian phraseology

The extensive amount of literature on phraseology written in Russian contains a large number of works discussing the semantic structure of PU (e.g. Dobrovol'skij 1998, Permyakov 1970, Melerovič 1998, Teliâ 1996, Solodub & Al'brecht 2003, Mel'čuk 1995, Kunin 1964, 1996, V. Žukov 1978 etc.). Each scholar's conception of this structure is initially influenced by the lexical-semantic framework which she/he adapted. In connection to this, there is an interesting comment made by A. Žukov (1999), who remarks that, as long as there is no universally accepted semaciological<sup>27</sup> conception, which, according to him, is unlikely in the foreseeable future, our ideas about the nature of both lexical and phraseological meanings, properties of their semantic structure etc. will only be hypothetical (A. Žukov 1999). The frameworks of structuralism and semiotics, which for a long time dominated Soviet linguistics, have left a clear mark on the methods of phraseological semantic studies, forcing them into a Procrustean bed of the taxonomic approach and componential analysis.

Thus, during the "classical period" of phraseology the prevailing method for the study of phraseological meaning was to compare it to the meaning of an equivalent unit of free syntax taken as a sum of lexical meanings of its components. This approach was used to determine the difference between the semantic shares of separate components in the overall meaning of the unit. Under these circumstances, the result always depended on the manipulations with the meanings of lexical components.

---

<sup>27</sup> Being a part of lexicology, semaciology studies the semantic structure of words as well as the semantic structure of the lexical system as a whole (Solodub & Al'brecht 2003: 48).



An example of such a method is phraseological application developed by V. Žukov (1978), which implies a comparison of a PU's meaning (in practice its dictionary definition) with the literal meaning of an equivalent free word-combination (if there is one available). By means of such comparison V. Žukov (ibid.) attempted to determine the degree of semantic unity of the lexical components and discover whether or not any component is used in its usual systemic meaning (literal or figurative). If all components are semantically deactualized (Rus. *деактуализованы* [*deaktualizóvany*]), the overall meaning of the PU is described with the help of words that do not and cannot form part of the PU as its lexical components.

For example, the meaning of the PU *ПЛЕВАТЬ В ПОТОЛОК* [*plevát' v potolók*] lit. 'to spit onto the ceiling' can be reproduced as 'to idle', 'to lounge', 'to loaf', 'to do absolutely nothing'. Since none of these words appear in the PU's lexical structure, V. Žukov (ibid.) draws a conclusion that lexemes constituting the phrase taken in its literal meaning are semantically incommensurable with the corresponding lexical components of this phrase taken as a PU. This is what he calls "an even deactualization of components" (Rus. *равномерная деактуализация компонентов* [*равnomérnaja deaktualizácija komponéntov*]) (ibid. 12-13). In the spirit of structuralism, V. Žukov (ibid. 10) understands lexical meaning to be "the inner and socially fixed conceptual content which is constituted by a set of differential (meaning-distinctive) and integral (meaning-combining) semantic features". However, he also admits that these features can be discovered in practice with the aid of dictionary definitions. This implies that instead of comparing the semantic structure one is comparing definitions: the dictionary definition of the PU is compared to the dictionary definitions of its lexical components. This method has been criticized by Teliâ (1996: 87-88) for the reductive treatment of meaning and ignoring its nominative and communicative aspects.

### 2.1.1 Phraseological meaning vs. lexical meaning

The Russian theory of phraseological semantics is originally based on the postulate of the fundamental difference between lexical and phraseological meaning. It has been put forward during the so-called "classical period" of Russian phraseology (1960-1970s) as one of the main arguments in favour of the separation of phraseology from the lexicon.

The primary task, pursued by phraseology scholars (Vinogradov and his school incl. Arhangel'skij, V. Žukov, Kunin, Šanskij, Popov and others) during this period, was to single out the object of phraseological studies into an independent linguistic discipline (Vinogradov 1946/1977).

In order to achieve this goal, it was necessary to demonstrate that phraseological units (PUs) are indeed special units of language system that considerably differ from both words and free word-combinations (i.e. free syntax) and even belong to a special phraseological level of language (Kunin 1964, 1996; Arhangel'skij 1964). Since this demonstration was attempted on both the syntactic and the semantic levels, phraseological units were ascribed both a specific structure and a specific meaning. By the middle 1960s phraseology had gained the status of an independent linguistic discipline, and by the 1980s phraseological semantics had firmly established itself as a field of phraseology devoted to the study of the specific meaning of PUs as a special type of language signs.

The term phraseological meaning together with argumentation for its necessity was coined in the 1960s by Kunin (1964) and Archangel'skij (1964) separately. The latter emphasized the status of phraseological meaning as a special linguistic category. At the same time, some scholars still continued using the term lexical meaning with respect to PUs claiming that, although PUs and words are not equivalent as language units, semantically they correspond to each other (Molotkov 1977: 29). Others (Šanskij 1963: 39) argued that the meaning of PUs equivalent to words or phrases bears a close analogy to word meaning.

V. Žukov (1978) shares Kunin's (1964) and Archangel'skij's (1964) opinion and argues that PUs have their own, specific meaning, different in many respects from the lexical meaning. Phraseological meaning is, as he puts it, "generalized and integrated", but the degree of integrity may be different. According to V. Žukov (1978: 20), the difference lies int. al. in the fact that the "semantic range" of PUs is broader than the semantic range of words, e.g. the Russian PU *с гулькин нос* [*s gúl'kin nos*] lit. 'the size of a dove's nose', id. 'less than nothing' does not simply mean 'little' but 'too little', i.e. this feature is expressed by the PU with a greater intensity than by the word. V. Žukov remarks that the meaning of a PU in most cases cannot be described by a single word, but a whole phrase. However, V. Žukov's argumentation does not seem to hold, since there is a large number of words that express intensity of state, activity or

character, e.g. *enormous* ‘very large indeed’, let alone the fact that phrasing word meaning can hardly be done in a way other than by using several words.

Teliâ (1996: 84) claims that “there is every reason to assert that meaning of idioms essentially differs from word meaning”. Like V. Žukov (1978), Teliâ (1996) too believes that the idiom meaning is much richer in detail than the word meaning – it has a larger semantic capacity. Teliâ’s (1996: 8) main claim is that the specificity of PUs as signs lies in the fact that they represent “microtexts” that include all kinds of information typical for representation of a situation in a text. She demonstrates this by the example of the Russian PU *ЗАДИРАТЬ НОС* [*zadirát’ nos*] lit. ‘to turn up one’s nose’, which does not merely mean ‘to be haughty, arrogant’ but also contains the implication that the person to whom this PU is applied was formerly socially equal with those to whom his attitude is expressed (Teliâ 1996: 89), as happens in Example (9) below. For this reason, the sentence in Example (10) below is not well-formed, since the Queen has not previously been equal with the members of parliament:

- (9) *Нашего сотрудника повысили, и он теперь нос задирает.*  
 [Náshego sotrudníka povysili i on tepér’ nos zadirájet.]  
 lit. ‘Our colleague has been promoted and now he **turns up his nose**.’
- (10) \**Королева Англии задирает нос перед парламентом.*  
 \*[Koroléva Anglii zadirájet nos pered parlámentom.]  
 \*lit. ‘The Queen of England **turns up her nose** at the parliament.’

Nevertheless, Teliâ’s (1996) argument can be contradicted, since the same meaning and implication can be expressed by the Russian verb *зазнаваться* [*zaznavát’s’a*] ‘to put on airs’ i.e. the meaning of this lexeme possesses the very quality (i.e. the ability to express pragmatic information) which Teliâ (1996) presents as a distinctive feature of phraseological meaning that sets it apart from lexical meaning.

### 2.1.2 Broad meaning

Several phraseology scholars point out that a certain indeterminacy of meaning is peculiar to many PUs. Consider Examples (11) and (12) below, featuring two different Finnish PUs:

- (11) *Pekka heitti lusikan nurkkaan.*  
 lit. ‘Pekka threw a spoon into a corner.’  
 id. ‘Pekka died.’
- (12) *Pekka heitti helmiä sioille.*

lit. 'Pekka threw pearls to pigs.'

id. 'Pekka caused a transfer of some entity, evaluated by the speaker as good, to a recipient, evaluated as inadequate in some way.'

Let us suppose that one's task is to interpret both of these sentences idiomatically. In (11) one does not need any additional context in order to tell what exactly Pekka did, since the Finnish PU *X HEITTÄÄ LUSIKKA NURKKAAN* | *NP<sub>SUBJ</sub>[N{X}] V{heittää 'cast'} NP<sub>OBJ</sub>[lusikka 'spoon'] PP<sub>ILL</sub>[NP<sub>SG</sub>{nurkka 'corner'}]* | lit. 'X casts a spoon into a corner, id. 'X dies' unambiguously denotes dying<sup>28</sup>. In (12) the picture is quite different. Without any context the exact nature of the action performed by Pekka remains unclear: he could perform Rachmaninoff's music, write a poem, discuss Plato's dialogues, offer financial or psychological support etc. At this rate, only contextual realizations of this PU would enable us to specify its meaning. The data collected from Usenet discussion groups which features the Finnish PU *X HEITTÄÄ HELMIÄ SIOILLE* | *NP<sub>SUBJ</sub>[N{X}] V{heittää 'throw'} NP<sub>OBJ</sub>[N<sub>PL</sub>{helmi 'pearl'}] PP<sub>ALL</sub>[N<sub>PL</sub>{sika 'pig'}]* | lit. 'X throws pearls to pigs', id. 'X causes a transfer of some entity, evaluated by the speaker as good, to a recipient, evaluated as inadequate in some way' exhibits considerable semantic variation. For instance, examples (13) and (14) below demonstrate two somewhat different meanings:

- (13) *siksi argumentointi-pyyntöt ovat pelkkää tyhjää jos toinen ei itse halua uskoa siihen, mihin toinenkin. siksi oman uskonsa argumentointi on sekä lapsellista, että typerää mutta ennen kaikkea tuota argumentoija heittää helmiään sioille [...].*<sup>29</sup>  
lit. 'That's why requests for argumentation are in vain if one party doesn't want to believe in what the other one does. That's why argumentation for one's own faith is both childish and stupid but above all the one who argues **throws his pearls to pigs** [...]'
- (14) *Sehän on siis aivan mielettömän hyvä mainossarja Ja lisäksi helmiä ei ole tarjoiltu sioille sillä tätä Hasan and Partnersin suunnittelemaa kampanjaa näytetään vain ja ainoastaan sivistyneille suomalaisille, ei ruotsalaiselle roskasakille joka nyt ostaa Arlaa kuitenkin.*<sup>30</sup>  
lit. 'So it is an extremely good series of ads. And besides, **pearls haven't been offered to pigs** because this campaign designed by Hasan and Partners is shown only to cultivated Finns, not to Swedish riffraff that buys

<sup>28</sup> It is worthwhile mentioning here that a few occurrences of this PU has been detected by me on the Internet meaning 'to give up' or 'to lose', but those were found exclusively in the context of sports-related discourse.

<sup>29</sup> GG: sfnet.keskustelu.filosofia/msg/e2eab3e751637056, Nov 24 1995

<sup>30</sup> GG: sfnet.huuhaa/msg/3723053eb53d1a86, Nov 6 1995

Arla anyway.'

In (13) the meaning can be paraphrased as 'to argue for one's faith with someone who does not want to believe in the same thing' and in (14) 'to show a good TV commercial to somebody who will buy the product anyway'. The task of a researcher is therefore to represent the indeterminate semantic structure of this PU so as to show how different semantic specifications are generated in the context.

The problem of semantic indeterminacy is dealt with in the linguistic literature in a very confusing way and suffers from terminological and descriptive divergences. Concepts like vagueness, ambiguity, generality, fuzziness etc. are often mixed up and their interpretation may differ (e.g. Zwicky & Sadock 1975, Kooij 1971, Lakoff 1970, Kempson 1977, Zhang 1998, Devos 2003). Generality is often understood as lack of specification, but the whole idea of generality is vague, i.e. there are no clear criteria that enable us to say whether an expression is general or not. Obviously, there are always conceptual specifications to be made (Devos 2003: 131). For example, if one states that *sister* is unspecified on the basis of the distinction between 'older sister' and 'younger sister' (Zwicky 1973: 100) one first has to prove that such specification is indeed necessary. If one begins to look for unnecessary specification<sup>31</sup> there is a danger of reducing the whole concept to pointlessness (Devos 2003: 131).

Several Russian linguists use the term broad meaning or eurysemy (Amosova 1963, A. Žukov & K. Žukov 2003; Alefirenko & Valjuh 1990). Amosova (1963), who has coined the term, gives the following definition of this phenomenon in lexical semantics:

By the broad meaning of a word we understand meaning which contains the highest possible degree of generalization, which becomes apparent in pure form only under isolation of the word from speech and which receives a certain narrowing and concretization when the given word is used in speech (Amosova 1963: 114).

Amosova (1963) points out that interpretation of the broad meaning is possible only in very general forms and is never equal and identical with the definition of its concretized variant in every particular case of this word's functioning. The broad meaning should not be confused with

---

<sup>31</sup> Cf. Lakoff's (1970: 357) statement that *kick* is unspecified in the sense that we do not know whether the action was carried out using the left or right foot, deliberately or accidentally.

polysemy. According to Amosova (1963), different meanings of a polysemous word co-exist within its semantic structure, but when the word is used in speech all meanings are eliminated but one. On the contrary, an eurysemous word is monosemantic outside context, but its single meaning has a very broad semantic reference, a possibility to point to an indefinite number of denotata. It contains in itself all possible concretizations caused by the context or the situation of speech. The context or the situation concretize, but neither change nor eliminate the broad meaning of the word, which remains as a base of any of its narrowed variants. Alefirenko & Valjuh (1990) and A. Žukov & K. Žukov (2003) talk about eurysemy in phraseological units. The latter aim to give a description of the semantic structure of eurysemous language signs by dividing it into a permanent part (core) and a variable part (periphery). The semes that constitute the peripheral part are finally determined in the context.

The definition given by Amosova (1963) is marred by same shortcomings as the above-mentioned approach to generality, i.e. it is not clear how the “highest possible degree of generalization” could be determined, and therefore there will always be uncertainty with regard to application of this term to a language sign. Moreover, without an explicit representation of the semantic structure any reasoning about its permanent and variable parts are practically pointless.

### **2.1.3 The connotative component in the multi-component model of PU semantics**

Several recent studies in phraseological semantics emphasize the multi-component nature of a PU's semantic structure (e.g. Solodub & Al'brecht 2003, Teliâ 1996). Solodub & Al'brecht (2003) find the same set of components within the lexical and phraseological semantic structures, namely:

- S (significative component): a conceptual basis, i.e. a conceptualized set of most significant properties, qualities of an object or situation, e.g. ‘a natural water body of considerable size enclosed by land’ for a *lake*;
- D (denotative component): in a broad sense, any fragment of reality and its idea to which the sign points, e.g. the interjection *y8bi* [uvý] ‘alas’ points to a certain mental state that the speaker wishes to express (sorrow, regret etc.);

- C (connotative component)<sup>32</sup>: “additional semantic or stylistic nuances which are put over the primary meaning of the word and serve to convey an emotive-expressive tint, imparting a tone of solemnity, naturalness, familiarity etc. to the expression” (Rosental and Telenkova 1985: 111);
- EC (ethnic-cultural component): when lexical or phraseological meaning reflects some realities or constructs specific for a certain nation or culture<sup>33</sup>, e.g. in the Russian PU *ЛЕЖАТЬ НА ПЕЧИ* [*ležát’ na pečí*] lit. ‘to lie on an oven’, id. ‘to do nothing, be idle’ such a component is *печь* [*peč’*] ‘oven, stove’ because it reflects the fact that Russian ovens were not only used for cooking and domestic heating, but also as a place for sleeping;
- STR (structural component) accounts for paradigmatic relations of language signs, their ability to fill certain semantic spaces, like LSG (lexical-semantic group), which unites lexical and phraseological units of the same class that have at least one integral seme, e.g. lexical and phraseological units denoting a state of idleness: *СИДЕТЬ СЛОЖА РУКИ* [*sidét’ složá ruki*] lit. ‘to sit with folded arms’, *ПЛЕВАТЬ В ПОТОЛОК* [*plevát’ v potolók*] lit. ‘to spit onto the ceiling’, *ЛЕЖАТЬ НА БОКУ* [*ležát’ na bokú*] lit. ‘to lie on one’s side’, *бездельничать* [*bezdel’ničát’*] ‘to idle’, *лентяйничать* [*lent’ájničát’*] ‘to be lazy’, *лодырничать* [*lódyrničát’*] ‘to loaf’.

According to Solodub & Al’brecht (2003), the difference between lexical and phraseological meaning lies in the obligatory vs. optional nature of the above-listed components. For lexical units S, D and STR are obligatory, and C and EC are optional. For PUs only EC is optional, while C is an obligatory and essential component. The conclusion drawn by the authors is that phraseological meaning differs from lexical in respect of the leading position which the connotative component always occupies in the structure of the former.

Teliâ (1990: 4, 32-46; 1996: 103-131) also talks about a macro-component model of meaning (both lexical and phraseological). This model should include all types of information manifested in the meaning

---

<sup>32</sup> One should keep in mind the fact that such concepts as signification, denotation and connotation lack a commonly accepted interpretation. The definitions of all components here are given as in Solodub & Al’brecht (2003).

<sup>33</sup> For a detailed account of cultural phenomena in PUs independent of Solodub & Al’brecht’s (2003) model, see Dobrovol’skij (1998) and Dobrovol’skij & Piirainen (2005).

of PUs. The form of this model is, as she puts it, declarative-procedural, i.e. it declares the content of components which implies the procedure of their processing. In concordance with the method of description conventionally adopted in phraseology, lexical meaning serves as a starting point and a background for Teliâ's (1990, 1996) explanation of the specific character of phraseological meaning. She divides lexical meaning into macro-components or information blocks of two types: descriptively oriented (information concerned with reference) and pragmatically oriented (expressing the speaker's/addressee's attitude towards the referential aspect of meaning). The former includes denotation<sup>34</sup>, which points to the standard idea (typical image) of the signified; the latter can be interpreted as connotative meanings. Like Solodub & Al'brecht (2003), Teliâ (1996: 89) seems to emphasize the role of connotation in the semantic structure of PUs, as she remarks that idioms are always loaded with the opinions and emotional attitude of the speaker.

Both above-mentioned models place a special emphasis on connotation in the structure of phraseological meaning. The problem is, however, that connotation is a very vague and ambiguous concept, which lacks generally accepted interpretation. An illustration in point is the fact that Teliâ (1996) defines connotation in a way that differs from that of Solodub & Al'brecht (2003). The latter accept Rosental & Telenkova's (1985: 111) definition of connotation as:

[...] additional semantic or stylistic nuances which are put over the primary meaning of the word and serve to convey an emotive-expressive tint, imparting a tone of solemnity, naturalness, familiarity etc. to the expression.

Teliâ (1996) understands by connotation virtually any pragmatically oriented component of the content plane of language units (morphemes, words, PUs and even segments of text), which:

[...] supplements their denotative and grammatical meaning on the basis of information correlated with pragmatic factors of different kind: with the associative background knowledge of language speakers about qualities or manifestations of the signified object or situation, with the rational-evaluative or emotive-evaluative attitude of the speaker towards the signified, with the stylistic registers which

---

<sup>34</sup> Teliâ (1990, 1996) interprets this term in a narrower sense than Solodub & Al'brecht (2003) and rejects the possibility of denotation for interjections and affective vocabulary in general.



characterize speech conditions or the field of language activity, social relations between participants of the speech situation, its forms etc. (Teliâ 1996: 107)

One should also distinguish connotation as a logical and philosophical concept from a linguistic and lexicographical one<sup>35</sup>. Furthermore, in linguistics this term is used very broadly to denote quite different phenomena, e.g. additional, accompanying components of word meaning, expressive and emotive overtones, modal and evaluative elements, pragmatic features, semantic associations, stylistic meaning etc. (Iordanskaja, Mel'čuk 1980: 192, Teliâ 1986: 5-6). There are also differences in approaches to connotation depending on whether it is regarded as personal associations or conventional, codified ones, or whether it lies outside meaning proper or is a rightful component of meaning (e.g. in the macro-component model).

In the Anglo-Saxon linguistic tradition the term is used mainly to refer to both socio-cultural and personal associations with a special emphasis on their individual subjectivity (Chandler 2002). In this tradition it is defined as "the vaguer associations of a word for a group or individual" (Cook 1992: 8), "a favourable or unfavourable evaluation by the speaker towards what they describe" (Partington 1998: 66), "attitudes of a society and of individuals", "the affective or emotional associations [...] which clearly need not be the same for all people who know and use the word" and "the personal aspect of meaning" (Kreidler 1998: 45). If one agrees to include personal subjective associations into the scope of connotation, one can argue that no language sign is purely denotative. It would also mean that all signs are truly polysemous in their connotative aspect of meaning, as there would virtually be as many connotations for each sign as there are language users: "The connotations which one person associates with a name may be different from the connotations which another person associates with the same name [...]" (Lyons 1977: 220). According to Kreidler (1998: 45) connotations vary from individual to individual but, because people do have common experiences, some words have shared connotations. Chandler (2002: 142) claims that connotations are not purely personal meanings; they are largely

---

<sup>35</sup> Definitions of connotation in logic and philosophy are not discussed here in order to avoid further confusion.

determined by codes. Partington (1998: 65–66) talks about three different types of connotation:

- Social or situational connotations that can indicate the class, origin, age or sex of the speaker, the relationship between the speaker and the listener and the formality of the register.
- Cultural connotations are a matter of associations which a lexical unit has within a culture; they are liable to modification as a society's values change.
- Expressive connotations imply a favourable or unfavourable evaluation by the speaker towards what they describe; they are highly personal and relatively more voluntary than social connotations.

A different view on connotation is expressed by Russian linguists, who would regard only widely recognized and conventionally encoded information as connotation, defining it as “inessential but stable features of the concept, which embody the evaluation of a respective object or a fact of reality adopted in a given language community” (Apres'jan 1974: 159); “elements of pragmatics that express cultural conceptions and traditions connected with the word, the practice of using a respective object prevailing in a given society” (Apres'jan 1995: 67); or the totality of associations fixed in a given society, which forms logical and emotive elements of the plane of content accompanying the lexical meaning and develops into a stereotype (Bartmiński 1980: 13-14). Connotation understood in a narrow sense excludes subjective associations:

Obviously connotations differ from other types of pragmatic information in a sense that they refer not to the individual user of the sign – the speaker, but to the language community. Thus the speaker who uses a lexeme that has certain connotations does not express his personal evaluation of the denoted object, as it was in the case of evaluation which is a part of pragmatic layer of the lexical meaning [...] (Kobozeva 2000: 92).

In the Russian linguistic tradition, connotation is often defined as a type of semantic-pragmatic information accompanying the meaning of a lexical unit and including additional semantic elements of certain types – expressive, stylistic, and evaluative. For example, the meaning of *mule* is ‘the animal which is the young of a donkey and a horse’ while its connotation in English is ‘stubborn’. Connotations are culture- and language-specific; for instance in Russian the word *мул* [*mul*] ‘mule’ lacks the connotation of stubbornness, which is instead ascribed to *осёл* [*os'ól*]

‘donkey’. Connotation manifests itself in a variety of language phenomena; among them are figurative meanings of words (cf. ‘a stubborn person’ of *mule*), simile (e.g. the Russian, the Finnish and the English synonymous expressions УПРЯМЫЙ КАК ОСЁЛ [*upr’ámyj kak os’ól*] lit. ‘stubborn as a donkey’, ITSEPÄINEN KUIN AASI lit. ‘stubborn as a donkey’ and AS STUBBORN AS A MULE), the meaning of derived words (e.g. the Russian ушач-умь [*išáčit’*] lit. ‘donkey-V-AFF’, id. ‘to work very hard’ or the Finnish pöllö-sti lit. ‘owl-ADV-AFF’, id. ‘stupidly’) and meanings of PUs (e.g. the Russian ЗАЯЧЬЯ ДУША [*zájačja dušá*] lit. ‘hare’s soul’, id. ‘a coward’ or the Finnish X<sub>ADE</sub> MENEE PUPU PÖKSYIHIN lit. ‘a bunny goes to X’s trousers’, id. ‘X becomes too frightened to do something’).

According to Solodub & Al’brecht (2003: 53-55) connotation has a hierarchical structure. They claim that it is formed by several subcomponents which are arranged in an algoristic order:

- emotive-evaluative
- expressive
- figurative
- word-formative

It is quite remarkable that Solodub & Al’brecht arrive at this set of components and the idea of their interconnection simply by juggling with some vague terms and definitions like the one of connotation as “additional semantic or stylistic nuances which are put over the primary meaning of the word and serve to convey an emotive-expressive tint, imparting a tone of solemnity, naturalness, familiarity etc. to the expression” (Rosental & Telenkova 1985: 111) and another of expressivity as “figurativeness, clearness, vividness of the mental content of speech” (Ahmanova 1968: 203-204). From these definitions Solodub & Al’brecht set off emotive-expressive and figurativeness in bold print and put them on different levels of their hierarchy, separating emotive from expressive without giving any explanation as to why it is so, except for remarking that “expressivity of the word is the very quality that contributes to the manifestation of its emotive-evaluative potential”, which is supposed to justify placing expressive on a lower level. In the same way one could just claim that evaluation depends on emotiveness and talk about emotive and evaluative as two separate levels instead of one. It is also obvious that the word-formative component is optional, as it is not necessary for a word to be derived with the help of a diminutive or pejorative suffix in order to have a connotation.

Iordanskaya & Mel'čuk (1980) give a quite narrow formal definition of connotation: "Lexical connotation of a lexical unit L is some characterization that L ascribes to its referent and which is not a part of its definition." If one applies this definition to *кляча* [kl'áča] 'jade', one arrives at the conclusion that since 'worn-out', 'broken-down', 'old' and 'useless' are parts of this word's definition they cannot be regarded as its connotation. Iordanskaya & Mel'čuk (1980) also offer two tests that allow distinguishing lexical meaning from connotation:

Test 1: A Lexeme L has a hypothetical connotation C. If adding an element with the meaning 'not C' to the lexeme L does not generate a contradiction, then C is truly a connotation. According to this test, 'stupidity' is a connotation of the Finnish word *pässi* meaning 'ram', because the phrase *fiksu pässi* 'clever ram' is not contradictory. This test also proves that 'old' is not a connotation of the Russian word *кляча* [kl'áča] 'jade', because *молодая кляча* [molodája kl'áča] 'young jade' is a nonsensical expression.

Test 2: A hypothetical connotation C names a function of an object, described by a lexeme L. C is an element of lexical meaning, if one can naturally infer from the fact that this object is out of order, that it fulfils the function C badly, otherwise C is a connotation. According to this test, 'to think' is a part of meaning for *head*, while 'to feel' is a connotation for *heart*.

A much broader definition is provided by Teliâ (1996: 107), who understands by connotation virtually any pragmatically oriented component of the content plane of different language units (morphemes, words, PUs and even segments of text), which supplements their denotative and grammatical meaning on the basis of information correlated with pragmatic factors of different kinds: with the associative background knowledge of language speakers about qualities or manifestations of the signified object or situation, with the rational-evaluative or emotive-evaluative attitude of the speaker towards the signified, with the stylistic registers which characterize speech conditions or the field of language activity, social relations between participants of the speech situation, its forms etc. The pragmatic nature of connotation lies in its subjective-modal nature: unlike denotation, it does not reflect an object or a phenomenon by itself, but rather the speaker's attitude or opinion towards it. Although Teliâ (1996: 107) remarks that the notion of pragmatics is broader than the connotative aspect of meaning, she tends

to believe that one can find as many connotations in the lexical meaning as there are pragmatic intentions expressed by it.

For Kobozeva (2000) an example of connotation is e.g. characteristics of 'stubbornness' and 'stupidity' of *осел* [os'ól] 'donkey', but not the negative evaluation of *кляча* [kl'áča] 'jade (a worn-out, broken-down, old or useless horse)', since she believes that a component that reflects the emotive-evaluative attitude of the speaker towards the object should be placed in the pragmatic layer of the lexical meaning. According to Kobozeva (2000), this pragmatic layer includes information about the speaker's attitude towards the denoted object<sup>36</sup> or towards the hearer, as well as about the pragmatic functions of the lexeme (i.e. the speech acts that the speaker can accomplish with the help of it). Although she admits that connotation belongs to the pragmatics of sign, she prefers to make a distinction between connotation in a narrow sense and other types of pragmatic information.

By and large, there is no consensus of opinion among linguists regarding the notion of connotation or its place in the lexical meaning. It has already been mentioned that e.g. Solodub & Al'brecht (2003) and Teliâ (1996) consider it a full-fledged component of meaning. Others would see connotation as "lying outside the core meaning" (Backhouse 1992: 297) or not being directly a part of semantics of the sign (Apres'jan 1995: 68). Another approach to description of connotation and denotation in relation to each other is adopted by Barthes (1972), who speaks about these in terms of orders of signification, denotation being the first order and connotation the second order of signification. Chandler (2002: 141) remarks that while some linguists may find it analytically useful to distinguish connotation from denotation, in practice such meanings are often difficult to separate. He quotes Voloshinov (1973: 105), who insisted that no strict division can be made between denotation and connotation since "referential meaning is moulded by evaluation [...] meaning is always permeated with value judgement". A similar opinion is expressed by Krongauz (2001: 139), who admits that a strict boundary between lexical meaning and connotation does not exist.

The last statement is especially applicable to the broad understanding of connotation (e.g. Teliâ 1996). For instance, the conceptual structure of

---

<sup>36</sup> Cf. the term modal frame coined by Wierzbicka (1967) to describe pragmatic elements of the sign.

the Finnish PU *X HEITTÄÄ HELMIÄ SIOILLE* | *NP<sub>SUBJ</sub>[N{X}] V{heittää 'throw'} NP<sub>OBJ</sub>[N<sub>PL</sub>{helmi 'pearl'}] PP<sub>ALL</sub>[N<sub>PL</sub>{sika 'pig'}]* | lit. 'X throws pearls to pigs', id. 'X causes a transfer of some entity, evaluated by the speaker as good, to a recipient, evaluated as inadequate in some way'<sup>37</sup> includes two propositions – one characterizing the contextual referent of *NP<sub>PL</sub>{helmi 'pearl'}* as GOOD, and another characterizing the referent of *NP<sub>PL</sub>{sika 'pig'}* as INADEQUATE (see Section 3.4.3.5 for details). These elements of the CS could be regarded as connotative (pragmatic) in Teliâ's (1996) sense, since they do not necessarily describe the objective characteristics of denotata (after all, the actual referents can in reality possess qualities that differ from the ones ascribed to them by the speaker), but rather the speaker's evaluation, i.e. his subjective attitude towards the participants in the referred situation. However, these characterizing elements cannot be regarded as connotation according to Iordanskaya & Mel'čuk (1980), who describe connotation as a characterization that PU ascribes to its referent, but which is not a part of its definitional meaning.

Another reason why the notion of connotation in its broad sense (i.e. encompassing any pragmatically oriented element, as understood by Teliâ 1996) can be regarded as problematic is because it would include cases where evaluation does not constitute a stable characterizing component of the PU's semantic structure (as in *HEITTÄÄ HELMIÄ SIOILLE* lit. 'throw pearls to pigs'), but can change depending on the situational context. A. Žukov (1999) claims that this "mobility of meaning", which reveals itself in "considerable fluctuations of the evaluative aspect", is a peculiarity of a number of idioms. He illustrates his point with the Russian PU *ОБВОДИТЬ ВОКРУГ ПАЛЬЦА* [*obvodít' vokrúg pál'ca*] lit. 'to turn smb. round one's finger', id. 'to dupe', which can change its evaluation from positive (15) to negative (16), and for the sake of convenience calls the phenomenon in question "appraisal eurysemy". However, in this case, one could easily question the status of evaluation as a component of meaning.

- (15) *Вот так уже шестой раз Гарри Поттер обвел вокруг пальца Темного Лорда!*<sup>38</sup>  
 [Vot tak užé šestój raz Gárrí Pótter obv'ól vokrúg pál'ca T'ómnogo Lórdá!]

<sup>37</sup> The detailed formal description of this PU and its construction family can be found in Chapters 3 and 5 of this book.

<sup>38</sup> <http://www.uralweb.ru/forums/messages.php?gid=24&id=9&page=35#t4707>, Oct 26 2008

lit. 'That's how Harry Potter **turned** the Dark Lord **round his finger** for the sixth time!'

id. 'That's how Harry Potter duped the Dark Lord for the sixth time!'

- (16) *Тринадцать минчан обвел вокруг пальца дуэт мошенников, который специализировался на квартирных сделках.*<sup>39</sup>  
*[Trinádcat' minčán obv'ól vokrúg pál'ca duét mošénnikov, kotóryj specializírovals'a na kvartírnyh sdélkah.]*

lit. 'A duo of swindlers, which specialized in real estate bargains, **has turned** thirteen residents of Minsk **round their finger**.'

id. 'A duo of swindlers, which specialized in real estate bargains, has duped thirteen residents of Minsk.'

All in all, the approach that postulates a particular importance of connotation in the structure of phraseological meaning (Solodub & Al'brecht 2003, Teliâ 1996) faces a major problem: the whole multi-component model of meaning, which operates with loosely defined semiotic notions (such as signification, denotation and connotation) as some sort of objectively existing and clear-cut macro-modules of linguistic meaning, is rather obscure. The lack of agreement amongst semioticians as to what connotation really is, as well as the absence of any coherent theory of linguistic meaning and methods of explicit formal analysis of semantic structure, result in the incapability to properly describe this component and show its place within this structure. But more importantly, by giving the status of the meaning component to a wide range of highly variable, subjective and context-dependent pragmatic information, this approach confuses competence with performance and semantic structure with contextual constraints.

## 2.2 Semantic representation of PUs in Meaning-Text Model (MTM) Theory

Dobrovol'skij & Piirainen (2005: 30) remark that Mel'čuk was the first to begin scientific research on phraseology in the framework of a consistent linguistic theory. This theory, called the Meaning-Text Theory (MTT) (e.g. Mel'čuk & Žolkovskij 1970, 1984, Mel'čuk 1973, 1974, 1981, 1995, 1996, 1998, 2001, 2004a, 2004b, 2007), indeed possesses an advantage over phraseological theory in that, unlike the latter, it attempts to construct a coherent model of natural language in general, and in fulfilling this task

---

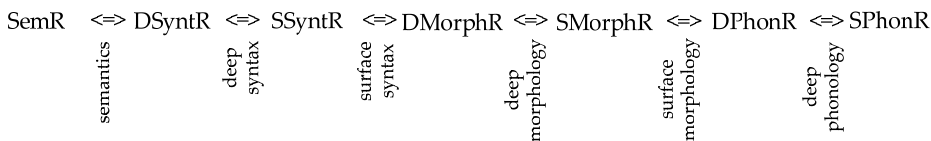
<sup>39</sup> [http://www.belarus.net/minsk\\_ev/99/russia/4\\_1/ni6.htm](http://www.belarus.net/minsk_ev/99/russia/4_1/ni6.htm), Apr 6 1999

makes use of formal methods of linguistic description. In this respect it is akin to Conceptual Semantics. A brief characterization of MTT can be found in Mel'čuk (1995) and a more comprehensive overview in e.g. Milićević (2006) or Wanner (2007). Here I will mainly concentrate on semantic aspects of this theory and Mel'čuk's approach to the semantic analysis of PUs.

MTT puts a strong emphasis on semantics, regarding language as a set of many-to-many correspondences between an infinite but denumerable set of meanings (a linguistic content that is communicated) and an infinite and equally denumerable set of texts (any fragment of speech). The Meaning-Text Model (MTM) of a natural language is considered to be a mapping of the form:

$$\{\text{Meaning}_i = \text{SemR}_i\} \Leftrightarrow \{\text{Text}_j = \text{PhonR}_j\} \mid 0 < i, j \leq \infty$$

Here SemR stands for Semantic Representation and PhonR for Phonological Representation. Between these two poles MTT assumes intermediate levels of representation, such as Syntactic Representation (SyntR) and Morphological Representation (MorphR). SyntR, MorphR and PhonR levels have two sub-levels, deep (D) and surface (S) representations, which results in a total of seven levels of representation (Figure 2 below):



*Figure 2 Levels of representation in MTM (Milićević 2006: 191)*

Mel'čuk (1995, 1998) also speaks about Conceptual Representation (ConceptR), which is "a mental reflection of perceived reality, of the speaker's encyclopedic knowledge relevant to the situation in question, of his intentions, preferences, wishes and goals, of his ideas about the addressee, etc", i.e. it contains virtually "everything that might be needed in order to say what the speaker wants to say about it" (Mel'čuk 1998: 25). According to Mel'čuk (1995), the speaker begins with a ConceptR and on the basis of it constructs the SemR according to the Concepts-Meaning Model (CMM), which associates elements and configurations of the ConceptR with elements and configurations of the corresponding SemR. Mel'čuk (1995: 172) calls this device (linguistic) pragmatics, while the



constructing of an appropriate PhonR for the SemR is regarded as language proper (Figure 3 below).

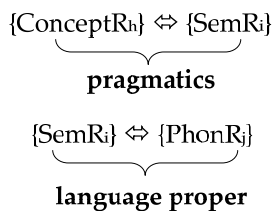


Figure 3 *Pragmatics vs. language proper in Mel'čuk's (1995) theory*

Although meaning-text correspondences are assumed to be bi-directional, natural language is nevertheless described by MTT predominantly in the direction of synthesis rather than analysis, (i.e. from meaning to text rather than from text to meaning), and from the viewpoint of production rather than comprehension. The study of synonymous linguistic expressions (or in MTT's terms, paraphrases) occupies the central place in MTT. Paraphrase appears to be one of the core concepts of MTT and is regarded as the main research tool in linguistics. Linguistic competence and communication is to a large extent understood in MTT as the ability to produce and understand paraphrases (e.g. speech production is referred to as virtual paraphrasing, i.e. choices between possible synonymous expressions of a starting linguistic meaning). Linguistic meaning of an utterance is defined as the invariant of paraphrases (i.e. as something, which is shared by synonymous utterances). Thus, the notion of the same meaning is taken to be prior to the notion of meaning, since it is regarded to be a simpler concept and "a primitive (intuitive) notion, underlying all our lexical knowledge" (Milićević 2006: 7). Primary importance is assigned to the lexicon. Lexical meaning is based on the notion of paraphrase as well: its analysis, largely relying on lexicographic definition of a lexical unit, involves semantic decomposition in terms of simpler meanings (Mel'čuk 1995, Milićević 2006). Thus, it is crucial for understanding MTT's approach to semantic description that its tools and methods are not initially designed to give an account of meanings as internalized mental representations or expressions of conceptual structure, but rather it is a model supposed to generate in parallel all synonymous syntactic variants of the same meaning.

The Semantic Representation (SemR) of a sentence includes four components: the Semantic Structure (SemS) the Semantic-Communicative

Structure (Sem-CommS), the Rhetorical Structure (RhetS) and the Referential Structure (RefS). Thus:

$$\text{SemR} = \langle \text{SemS}, \text{Sem-CommS}, \text{RhetS}, \text{RefS} \rangle$$

The Sem-CommS describes all the properties of a sentence which have to do with its communicative aspect (Thematicity, Focus, Perspective, Givenness, etc.); the RhetS, encodes the artistic, or aesthetic, intentions of the speaker (does he/she want his/her utterance to be neutral, ironic, pathetic, humorous, etc.) and his/her stylistic choices; and the RefS specifies the referential status and the concrete real-world referents for semantic configurations. The central (or carrying) component of the SemR is the SemS. This notion is used in MTM to denote propositional, or situational, meaning of synonymous utterances. The SemS mirrors a situation in the real or an imaginary world (including the speaker, his/her opinions, feelings, intentions, etc.) as it is reflected in speech, hence the name situational meaning. This meaning is expressed by logical propositions, hence the name 'propositional meaning' (Mel'čuk 2001: 4).

Formally, the SemS is represented in a form of a semantic network, which is used to symbolize decomposition of propositional meaning into semantemes (nodes of the network) and predicate-to-argument relations between them (connectors of the network). There are two kinds of semantemes: functors, which are predicates, quantifiers and logical operators, and names, which are names of (classes of) objects, including proper names (Mel'čuk 2001: 5). The connectors are labelled with numbers specifying predicate-to-argument relations: e.g. in the sentence *John sees Mary*, 'John' and 'Mary' are semantic arguments of the predicate 'see'. This can be represented either in the form '*see*'('John'; 'Mary') or in the form '*John*' $\leftarrow$ 1-'*see*'-2 $\rightarrow$  '*Mary*'. The arrows indicate semantic dependencies, while the numbers distinguish different arguments of the same predicate (Mel'čuk 2004a). Semantemes reflect the intuitions in paraphrasing complex lexical meanings with locutions involving simpler lexical meanings. Thus, the SemS is not presumed to be universal; it is fully language-specific, since even the simplest lexical and grammatical meanings of languages tend to differ (Kittredge et al. 1988). Semantemes themselves can be complex meanings, semantically decomposable to simpler meanings. The major problem here is that such decomposition relies on a lexicographic definition of a lexical unit having this semanteme as its signified. Thus, semantic description is equated with lexicographic definition (Milićević 2006: 8, 28-29), which by itself cannot

be regarded as being composed of semantic primitives<sup>40</sup>. And even if it was, the very idea that lexical meaning can be exhaustively decomposed into a finite set of semantic primitives, which is taken in MTT for granted, is quite debatable (see e.g. Jackendoff 1999: 112f; 2003: 334f for more discussion on this matter).

Mel'čuk (1995, 1998) introduces several auxiliary notions which need to be mentioned:

- A linguistic sign  $X$  as a triplet  $X = \langle 'X'; /X/; \Sigma_x \rangle$ , where  $'X'$  is the signified,  $/X/$  the signifier, and  $\Sigma_x$  the syntactics of the sign (i.e. the constraints on the co-occurrence of  $X$  with other signs).
- The operation of linguistic union  $\oplus$  unites linguistic items according to general rules of language and to syntactics of these items, while constructing expressions of higher order. It is similar to arithmetical summation, but linguistic union is much more complex than simple addition.  $X \oplus Y$  denotes the regular union of signs  $X$  and  $Y$ ,  $'X' \oplus 'Y'$  is the regular union of signifieds  $'X'$  and  $'Y'$ , and so on.
- The phrase  $A \oplus B$  composed of lexemes  $A$  and  $B$  is “free” iff it meets two conditions:

Condition 1: Its signified  $'X' = 'A \oplus B'$  is “unrestrictedly” (i.e. using any of the applicable lexical selection rules of a given language) and “regularly” (i.e. according to general combination rules of language) constructed on the basis of the given ConceptR out of the signifieds  $'A'$  and  $'B'$  of the lexemes  $A$  and  $B$ . Thus  $'A \oplus B'$  is a regular sum of  $'A'$  and  $'B'$ . Obviously, the signified is equal to SemR.

Condition 2: Its signifier  $/X/ = /A \oplus B/$  is unrestrictedly and regularly constructed on the basis of the SemR  $'A \oplus B'$  out of the signifiers  $/A/$  and  $/B/$  of the lexemes  $A$  and  $B$ . Thus  $/A \oplus B/$  is a regular sum of  $/A/$  and  $/B/$ .

If any or both of these two conditions are violated, the phrase  $AB$  is not free, and therefore represents a PU (or a phraseme in Mel'čuk's terminology).

Now, let us turn to the methods of PU analysis within the framework of MTT. Mel'čuk's (1995: 176-179; 1998: 28-31) classification of PUs is based on distinctions depending on which of the above conditions for a

---

<sup>40</sup> Jackendoff (2003: 335f) quotes Fodor et al. (1980), who argue that lexical meanings cannot be constructed by combining other lexical meanings, using the principles that also combine words into phrases. Jackendoff's (2003) answer to their argument is that there can be non-definitional forms of decomposition to the sublexical units that cannot individually serve as word meanings and whose combination principles are not the same as the principles of phrasal word combination.

free phrase are violated, and includes two classes, the second one including three sub-classes:

1. Pragmatic phrasemes (pragmatemes) are expressions where either of the two following violations are true:
  - a. Conditions 1 and 2 are violated in the sense that the signifier and the signified are regular, but not unrestrictedly constructed. For example, according to convention, food packages in Finland can have either of two inscriptions: *PARASTA ENNEN* ... 'best before ...' for foods with a longer shelf-life, and *VIIMEINEN KÄYTTÖPÄIVÄ* ... 'the last day of usage ...' for perishables, but inscriptions like *#ON NAUTITTAVA ENNEN...* 'to be consumed before ...' or *#ÄLÄ KÄYTÄ ... JÄLKEEN* 'do not use after ...' would be pragmatically inappropriate in this context.
  - b. Condition 1 is violated but Condition 2 is not, i.e. for the given ConceptR only the given signified ' $A \oplus B$ ' is possible, but it can be expressed unrestrictedly. As an example of this type of phraseme Mel'čuk mentions signs in United States libraries meant to prohibit talking, which can include *NO TALKING PLEASE*, *PLEASE DO NOT TALK*, or *PLEASE BE QUIET*, but not *#DON'T MAKE NOISE PLEASE*, *#PLEASE DON'T SPEAK WITH EACH OTHER* or *#KEEP SILENT PLEASE*.
2. Semantic phrasemes, where Condition 1 is not violated, in the sense that the signified ' $X$ ' is constructed unrestrictedly, although not regularly<sup>41</sup>, but Condition 2 is, i.e. the signifier  $/X/$  is not constructed unrestrictedly. This violation can come about in three ways:
  - a. Idioms (full phrasemes), instead of the expected regular sum ' $A \oplus B$ ' of the signifieds ' $A$ ' and ' $B$ ', have a signified ' $C$ ', which includes neither ' $A$ ' nor ' $B$ '. For example, *TO SHOOT THE BREEZE*, *TO SPILL THE BEANS*, *TO PULL X'S LEG*. Idioms are formally described as  $AB = \langle 'C'; /A \oplus B/ \rangle \mid 'C' \not\subset 'A' \ \& \ 'C' \not\subset 'B'$ .<sup>42</sup>
  - b. Quasi-idioms (quasi-phrasemes), e.g. *TO GIVE THE BREAST TO X*, *TO START A FAMILY*, *BACON AND EGGS*. Here the signified of  $AB$  includes the signifieds of both constituent lexemes, but also contains an unpredictable addition ' $C$ ':  $AB = \langle 'A \oplus B \oplus C'; /A \oplus B/ \rangle \mid 'C' \neq 'A' \ \& \ 'C' \neq 'B'$ .
  - c. Collocations (semi-phrasemes), e.g. *TO LAND A JOB*, *TO LAUNCH AN ATTACK*, *TO STAND COMPARISON*, are formally described as  $AB = \langle 'S';$

<sup>41</sup> I am having a hard time understanding how the condition which demands that a signified is both unrestrictedly and regularly constructed can be regarded as not violated if the signified is not regularly constructed.

<sup>42</sup> Mel'čuk (1995: 177) does not include syntactic descriptions in these formalizations as they would be "irrelevant in this context".

$/A/ \oplus /B/; \Sigma_{AB}>$ .<sup>43</sup> In order to be classified as a collocation, an expression has to satisfy all of the following conditions:

- i.  $AB$ 's signified ' $S$ ' includes the signified of one of its components, for instance, of  $A$  as its semantic dominant: ' $S \supset \underline{A}$ '.
- ii.  $A$  is selected by the speaker for its signified ' $A$ ' independently of  $B$  and unrestrictedly.
- iii.  $B$  is not selected unrestrictedly for its signified ' $S$ ' (which is the difference ' $S - A$ '):  $B$  is selected as a function of  $A$ .

Mel'čuk's tool for description of restricted lexical co-occurrence in collocations (as well as semantic derivatives, which will not be discussed here) is Lexical Functions (LF). The noun function in this term is used in its mathematical sense  $f(x) = y$ , i.e.  $Function(Argument) = Value$ . The adjective lexical indicates that  $f$ 's domain of definition and the range of  $f$ 's values are both lexical expressions (1998: 31-32; 2007: 121). A Lexical Function  $f$  is a function that associates with a given lexical unit  $L$ , which is the argument, or keyword, of  $f$ , a set  $\{L_i\}$  of (more or less) synonymous lexical items – the value of  $f$  – that are selected contingent on  $L$  to express a specific meaning corresponding to  $f$ . Thus  $f(L) = \{L_i\}$ . In other words, an LF is a “very general and abstract meaning that can be expressed in a large variety of ways depending on the lexical unit to which this meaning applies” (1995: 186). According to Mel'čuk, about 60 simple standard LFs have been empirically established. Among them is, for instance, the LF **Magn**, which stands for an intensifier ‘intense(ly)’, ‘very’:

**Magn**(*yhteistyö* ‘cooperation’) = *tiivis* ‘1 serried 2 concise 3 compact 4 taut 5, 6 compact 7 heavy, compressed 8, 9 dense 10 tight 11 close, intensive’<sup>44</sup>

**Magn**(*kilpailu* ‘competition’) = *kireä* ‘1 charged 2 tight-fitting 3 tense 4 rigid, unrelaxed, stringent 5 harassed, uptight, wired 6 edgy 7 severe 8 stiff 9 compact 10 compressed 11 constricted 12 difficult, tricky 13 tight, strained, taut 14 on edge’

In Finnish, the adjectives *tiivis* and *kireä* work as intensifiers in collocations *TIIVIS YHTEISTYÖ* ‘close cooperation’ and *KIREÄ KILPAILU* ‘close

<sup>43</sup> This formula appears in Mel'čuk (2007). However, earlier in Mel'čuk (1995) and Mel'čuk (1998) a quite different formal description of a collocation is presented:  $AB = \langle 'A \oplus C'; /A \oplus B/ > \mid 'C'$  is expressed by  $B$  such that  $/A \oplus B/$  is not constructed unrestrictedly.

<sup>44</sup> Here, lexical indices in definitions are presented according to the MOT Finnish-English 4.8 Finnish-English-Finnish dictionary.

competition'<sup>45</sup>, while in English the same LF applied to the lexical units *cooperation* and *competition* is expressed by the adjective *close*.

For a more comprehensive account on LFs see e.g. Mel'čuk (1995, 1998 and 2007). The advantages of LFs in accounting for combinatorial abilities of lexical units have become widely recognized by many collocation researches. However, being a powerful descriptive tool, LFs are not free from disadvantages: although the number of LFs is limited, there is, in principle, nothing that prevents introducing additional ones (Fontenelle 1998 actually does so). Thus, the notion of LF becomes blurred, especially taking into consideration the fact that the definition of LF presented by Mel'čuk (1995: 187) also includes cases where a function *f* is applicable to one lexical unit *L* only. As Fontenelle (1998: 202) points out, in some cases assignment of an LF to a pair of collocations can be difficult due to the problem of choosing among several possible alternatives, while no battery of tests is available to ensure consistency in the assignment of LFs.

As long as collocations lie outside the scope of the present study, LFs will not be discussed here in more detail. However, they have to be mentioned, since Mel'čuk also attempts to apply them in his analysis of another class of PUs, namely idioms (full phrasemes). Mel'čuk (1995: 205f) turns to the problem of syntactic transformations of PUs and argues that the problem of applicability/non-applicability of certain syntactic transformations to a particular PU is irrelevant and disappears if one describes the PU's meaning "in a rigorous enough way". He illustrated his point with two examples: (TO) KICK THE BUCKET '(to) die' and (TO) BREAK Y'S HEART '(to) make Y feel very sad and/or hopeless'. What one has to explain is why the former cannot be relativized or passivized, while the latter can. The difference in their syntactic behaviour, claims Mel'čuk (1995: 206), follows from the differences in their meaning: "it is necessary and sufficient to describe the meaning of these expressions properly – and then no special syntactic indications concerning applicable transformations are needed for them". However, Mel'čuk himself does not provide us with any explicit semantic analysis of the given expressions. His examples of idiom representation in MTM are in fact

---

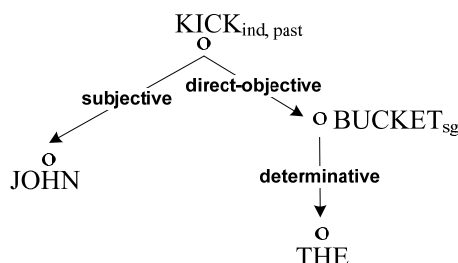
<sup>45</sup> The variant *TIIVIS KILPAILU* is also possible, but far less common than *KIREÄ KILPAILU*. An advanced syntax query in Kielipankki [bf="kireä"] [bf="kilpailu"] returns 55 hits compared to 2 hits for [bf="tiivis"] [bf="kilpailu"]. The collocation #*KIREÄ YHTEISTYÖ* is not possible. The query [bf="kireä"] [bf="yhteistyö"] gave no hits.

syntax-oriented. Idioms are represented in MTM in two ways (Mel'čuk 1995: 218):

1. On the DSynt (deep syntax) level an idiom is represented as a single node. Thus, the sentence *John kicked the bucket* has the form:

I  
JOHN o←o [KICK THE BUCKET]<sub>ind, past</sub>

2. On the SSynt (surface syntax) level an idiom is represented as an SSynt-tree, containing “normal” lexemes (*kick*, *the*, *bucket*). This is done according to “indications stored in its entry in the lexicon”:



As for the SemR, Mel'čuk's (1995) analysis does not go beyond simple paraphrasing. The SemR for *Pete kicked the bucket* is 'Pete died'. As for the expression *Pete broke Mary's heart*, Mel'čuk (1995: 206) postulates the following: the speaker starts with the “idea” ‘Pete made Mary feel very sad and/or hopeless’, but, since *heart2* is defined in LDCE as ‘Y’s imaginary organ of feelings’, “he can use this meaning to verbalize his idea as ‘Pete caused that Mary’s imaginary organ of feelings senses utter sadness and/or hopelessness’”. The latter paraphrase is, according to Mel'čuk (ibid.), the SemR of this expression, which is not an idiom after all: it is “a collocation, where a separate lexeme *heart2* co-occurs with the value of the LF **CausFact**<sup>46</sup>, given the constraint that the feeling sensed by Y is ‘utter sadness and/or hopelessness’”. However, since the lexical meaning of *heart2* does not include a restriction on the character of feelings, by imposing such a constraint on the semantic dominant of the supposed collocation, Mel'čuk (1995) violates his own definitional condition of collocation, according to which the base of the collocation has to be selected for its signified independently of its collocate and

<sup>46</sup> This complex LF is a combination of two simple LFs – **Caus** ‘cause [do something so that situation begins occurring]’ and **Fact<sub>oi</sub>** ‘fulfill the requirement of L’ [‘do with L what you are supposed to do with L’ or ‘L does with you what L is supposed to do with you’].

unrestrictedly (Mel'čuk 2007: 120). Mel'čuk (1995: 208-213) uses the following logic:

- Transformations like passivization, relativization and clefting are semantically driven.
- Full idioms are by definition non-decomposable and therefore cannot undergo semantically driven transformation.
- Multilexemic expressions which admit such transformations are not idioms; if at least one of the expression's parts is accessible to (almost) all semantically driven transformations it **MUST** be dissolved to separate lexemic parts and represented as a collocation (the Principle of Semantic Accessibility).

A similar claim is made by Mel'čuk (1995: 207) about the idiom *(TO) PULL STRINGS* id. '(to) use personal contacts among people in charge in order to obtain something that cannot be obtained otherwise' in order to justify its passivization: it is also a collocation where *pull* is a value of the LF **Real**<sub>1</sub><sup>47</sup> (*strings*), while *strings* is a separate lexeme meaning 'personal contacts among people in charge, which may be used in order to obtain something that cannot be obtained otherwise'. If this was indeed so, the same lexeme with the same meaning could be unrestrictedly selected by other expressions as well. However, the test in (17) and (18) below seems to show the opposite:

(17) *As Slovakia is a small country **personal contacts** play a key role. And the best way of finding a job is through **personal contacts**.*<sup>48</sup>

(18) *?As Slovakia is a small country **strings** play a key role. And the best way of finding a job is through **strings**.*

Mel'čuk (1995) himself admits this peculiarity and suggests the following solution: *strings* is a unique lexeme, i.e. a lexeme that co-occurs with only one other lexeme. However, he understands that such a notion would not prevent one from arbitrary postulation of unique lexemes in any idiom if it produces some advantage for its description, and admits the need for constraints. As one such constraint he proposes the Principle of Regular Polysemy, according to which a part *A* of an expression *E*' = *A*'*B*' can be singled out as a unique lexeme and this phrase can be dissolved into separate lexemic parts *A*' and *B*' if the following three conditions are met:

---

<sup>47</sup> The LF **Real**<sub>0/i</sub> is synonymous with **Fact**<sub>0/i</sub>.

<sup>48</sup> [http://www.eurograduate.com/planning\\_overview.asp?eid=4&id=37](http://www.eurograduate.com/planning_overview.asp?eid=4&id=37), Jan 3 2010



1. There is another phrase  $E = AB$  homophonous with  $E'$ , its lexical constituents  $A$  and  $B$  being homophonous with constituents  $A'$  and  $B'$  of  $E'$ .
2.  $E'$  and  $E$  stand in a regular polysemy relation.
3.  $A'$  and  $A$  stand in a regular polysemy relation.

But how circumscriptive are these conditions after all? The fact that the Principle of Semantic Accessibility is a prescription, while the Principle of Regular Polysemy is only a permission, indicates that the former principle entails the latter, but not vice versa. Some idioms, remarks Mel'čuk, have in their meaning an appropriate component that "could be picked out for passivization, relativization etc.", but which does not, however, correspond to a separate lexeme. For instance, the lexeme *beans* in the idiom (TO) *SPILL THE BEANS* could mean 'a secret supposed not to be revealed', but it does not. How do we know it? Mel'čuk's (1995) answer to this question is: "we know this because of the impossibility of \**The beans were spilled by X*, \**The beans that X spilled*, etc.", i.e. of semantically driven relativization and passivization transformations. If it happens that this idiom starts admitting these and other <sup>49</sup> semantically driven transformations (or de facto already started, which is demonstrated in (19) – (22) below), then *beans* is a separate lexeme meaning 'a secret supposed not to be revealed':

- (19) *A: Interviews in the UK mags have engineers saying "I use certain FX" whereas in the US mags the pros<sup>[sic]</sup> interviewed are more prone to 'spill the beans' on techniques, often detailing exactly how they did stuff. UK mags let the the<sup>[sic]</sup> engineers get away with guarding their 'secrets' so are of less use to those wanting to pick up new tricks.*  
*B: Yeah.... but an awful lot of those **beans that are spilled** aren't exactly truthful ones. I'd rather be told nothing than be told something wrong.<sup>50</sup>*
- (20) *A: IIRC he was tried, convicted and sent to jail (just as was to be expected). Rumours of plea bargaining to stop him spilling the beans but no proof (of course!)*  
*B: Of course ... but, has anybody heard any **beans that he did spill**?<sup>51</sup>*
- (21) *When a child was born in 1934 - Maria Pia - the crown princess' gynecologist revealed the royal secret while lecturing at the University of Naples on his*

---

<sup>49</sup> It still remains unclear whether the idiom has to admit all semantically driven transformations (Mel'čuk 1995: 207) or "almost all" as is formulated in the Principle of Semantic Accessibility (Mel'čuk 1995: 210). In the latter case, one would also need to establish what amount of transformations qualifies as "almost all".

<sup>50</sup> GG: rec.audio.pro/msg/f1a98843abd9af96, Dec 23 2000

<sup>51</sup> GG: demon.local/msg/51d2b7075ac1fd8e, May 16 1996

*experiments with artificial insemination. {This means the story was already known as early as 1934!!!! **The beans were spilled by Marie Jose's own doctor.**}*<sup>52</sup>

(22) *For Xara fans, do be aware that Xara 2.0 is imminent. **The beans were spilled by Gary Priester** in his March tutorial.*<sup>53</sup>

But how shall one deal with the fact that no English dictionary contains the lexeme *beans* in the sense 'a secret supposed not to be revealed', but still one can find examples like in (19) – (22) above? Does it mean that, as Mel'čuk (1995) suggests, for the speakers for whom these transformations are valid, *beans* is a unique lexeme, while for those who reject them it is not? Mel'čuk (1995) himself accepts only a restricted passivization *THE BEANS WERE SPILLED* without the agent phrase *by X*, but he does not see *beans* here as a separate lexeme. Such an approach makes the notion of the unique lexeme somewhat subjective. Obviously, one cannot possibly know whether the authors of (19) – (22) above do or do not regard *beans* as a unique lexeme.

Adherence to the Principle of Semantic Accessibility forces one to admit unique lexemes, but why do one has to follow this principle in the first place? It seems that Mel'čuk (1995) introduces it only because it provides advantage for his own description, i.e. justifies his treatment of certain idioms as collocations. As I have already mentioned earlier, MTT with its orientation on synthesis and dominant role of paraphrase (synonymy) does not have a theory of meaning as such, but rather a theory of the same meaning. The same holds for its descriptive methods: LFs are suited for analysis of cases of restricted lexical co-occurrence, but their applicability to the semantic analysis of idioms is doubtful. In the absence of any other elaborate methods of formal semantic analysis than LFs, Mel'čuk is eager to impose this descriptive tool on idioms by representing them as collocations.

The problem of variation in PUs, addressed by Mel'čuk (1995), can be dealt with in a completely different manner, i.e. without any necessity to attribute independent meanings to PU components. In Chapter 3 Section 3.4.3.3, I will argue that analyzability of PUs can be approached in terms of referentiality of their syntactic constituents, and in Chapter 5 this approach will be applied to variation in PUs.

---

<sup>52</sup> GG: alt.talk.royalty/msg/303a0543d1fdf1c7, Mar 9 1999

<sup>53</sup> GG: comp.graphics.apps.corel/msg/4bf1c437138acdc9, Mar 3 1998

## 2.3 Semantic representation of PUs in cognitive-linguistic models

### 2.3.1 Cognitive Grammar

Cognitive Grammar developed by Langacker (1987: 57) understands grammar as a structured inventory of conventional linguistic units. Langacker (1987: 77) defines linguistic units as bipolar symbolic units that consist of a semantic and a phonological pole. Form and meaning are linked through symbolization, i.e. the entrenched association between the semantic and phonological structures. The distinction between grammatical modules (lexicon, morphology, syntax) is abolished and replaced by a continuum of symbolic structures (Langacker 1987: 54). On this continuum traditional syntactic, morphological or lexical structures constitute form-meaning pairings with different degrees of complexity and schematicity (Langacker 1991: 16).

Figure 4 below is Langacker's (1987: 66) representation of a symbolic structure. It consists of a conventional linguistic unit (the sanctioning structure), which sanctions the target structure - a particular usage event, i.e. "a symbolic expression assembled by a speaker in a particular set of circumstances for a particular purpose". The target structure as such is not directly given by the grammar of a language and its non-unit status is indicated by rectangles with curved corners.

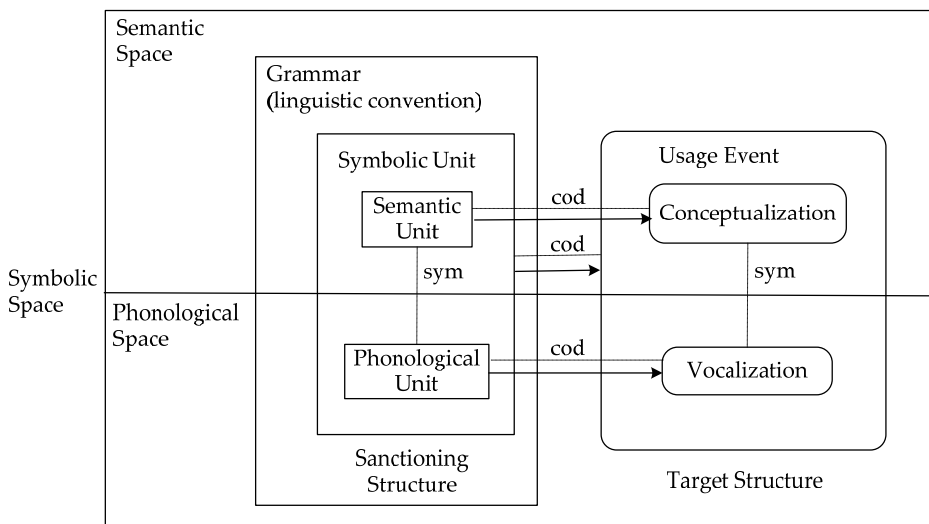


Figure 4 Langacker's (1987: 77) representation of a symbolic structure

Sanction is defined as “the motivation afforded a novel structure by the conventional units of a grammar” (Langacker 1987: 492), although in practice it is reduced to categorization: a conventional unit defines a category and sanctions a target category to the extent that the latter is judged by the speaker to be a member of the category (ibid. 68). Categorization in turn depends on the relation of schematicity: the sanctioning structure bears the relation of schematicity to the target structure. Schematicity can be equated with the relation between a superordinate node and a subordinate node in a taxonomic hierarchy. Langacker calls the superordinate structure a schema, and the subordinate structure an elaboration or instantiation of the schema (ibid. 68). Structure A is a schema with respect to structure B when A is compatible with the specifications of B but characterizes corresponding entities with less precision and detail (ibid. 492). Relation of schematicity in Figure 4 above is represented with a solid arrow. Conceptualization is an instantiation of a semantic unit and vocalization of a phonological unit, and the whole target structure is an instantiation of the sanctioning structure. Symbolization (sym) is the relation between a structure in semantic space and one in phonological space, e.g. in the target structure vocalization (i.e. actual articulation) symbolizes a detailed, context-dependent conceptualization. Coding (cod), defined as the task of finding appropriate linguistic expression for a conceptualization (ibid. 65), takes place across the boundary between convention and usage. Both symbolization and coding depend on correspondences represented with a dotted line (ibid. 77).

Langacker (ibid. 66, 68) claims that sanction is a matter of degree and speaker judgment. He distinguishes between full schematicity (full sanction) and partial schematicity (partial sanction). Full sanction involves an elaborative relationship between the sanctioning unit and the target structure. It occurs when an instantiation (the target) is fully compatible with the specifications of its schema (the sanctioning unit), but is characterized in more detail. If there is some conflict between the specifications of the sanctioning and the target structures, their relation is only of partial sanction and involves extension. Langacker (ibid. 69) also equates partial sanction with deviance or ill-formedness.

Langacker (ibid. 92-93) applies this theoretical model to figurative language. According to him, the latter is also a type of categorization, which involves partial schematicity and results in a bipartite

conceptualization including a literal sense as a sanctioning structure (SS) and a figurative sense as a target structure (TS) (Figure 5 below). The entire categorizing structure has a unit status.

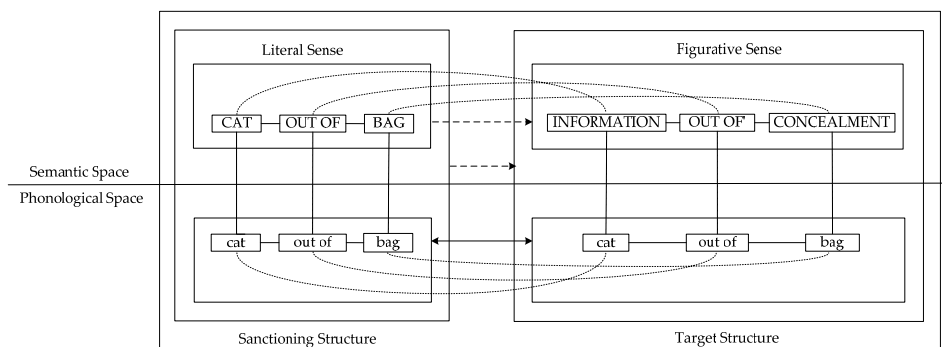


Figure 5 Langacker's (1987: 93) representation of "the essential components and relations" within the PU *THE CAT ... OUT OF THE BAG*

Figure 5 above is Langacker's (1987: 93) analysis of the English PU *the cat ... out of the bag* (Langacker himself uses this form in order to simultaneously account for two different constructions – *LET THE CAT OUT OF THE BAG* and *THE CAT IS OUT OF THE BAG*). At the phonological pole a solid double-headed arrow stands for the relationship of identity, which is further resolvable into identity relations between the individual phonological components of SS and TS. Partial schematicity relation between the sanctioning and the target structures is indicated by a broken-line arrow. It is presented as a global relationship at the semantic pole (semantic extension) and is also decomposed into local correspondence relations between semantic components: [INFORMATION] is a semantic extension of [CAT], [OUT-OF']<sup>54</sup> – of [OUT-OF] and [CONCEALMENT] – of [BAG] (ibid. 94). The status of these components remains unclear: are they some language concepts, and, if so, concepts of what? They are not defined and unanalyzable.

According to Langacker (ibid. 69, 157), "categorization based on partial sanction is the kind described in the prototype model, where a category is defined in terms of prototypical instances". The target structure is a "semantic extension from the prototype", based on the speaker's perception of similarity or association between the original (sanctioning)

<sup>54</sup> An apostrophe symbolizes a variant which is more specific than the corresponding unit, e.g. structure X' is a variant of X (Langacker 1987: 68).

sense of an expression and its extended sense. It means that in order to be able to draw a direct correspondence link from [CAT] to [INFORMATION], there has to be some similarity or association between them. But as Langacker himself admits, the conflict in specifications is blatant, and it is hard to perceive any similarities between these semantic components. Despite this conflict, Langacker (ibid. 93) claims that *cat* is attributed a meaning roughly equivalent to ‘information’, while *bag* conveys notion akin to ‘concealment’. Finally, he states that the integration<sup>55</sup> of semantic units is parallel in the sanctioning and the target structures.

In what follows, I will test the correctness of Langacker’s (1987) semantic analysis of the given PU by turning to the actual examples of this PU’s variation, such as modification and lexical substitution. This will be done primarily because Langlotz (2006), who in his study of idiom variation<sup>56</sup> relies on Langacker’s theory, claims that topic-indicating lexical substitution and premodification work as proof for the semantic analyzability of isomorphic PUs. If this is true, and if Langacker’s (1987: 94) analysis in Figure 5 above is correct, then one needs an explanation as to why, despite such a blatant conflict in specifications between the component [CAT] and its target, the component tends to undergo topic-indicating lexical substitution and premodification, while the component [BAG], which “bears a natural and salient relation to [CONSEALMENT]”, is very reluctant to such variation.

According to Langlotz (2006: 268), in analyzable<sup>57</sup> PUs, if an adjectival premodifier is interpreted as an internal modifier, it defines a subclass of the phrase-induced figurative sense of the idiomatic head-noun. In other words, if the phrase-induced figurative sense of *cat* is indeed ‘information’, as is stated by Langacker (1987), the premodifier will function to qualify this particular sense, i.e. *renewable cat* in (23) below must be paraphrased as *renewable information* and *nuclear cat* in (24) below as *nuclear information*:

---

<sup>55</sup> Langacker (1987: 490) defines integration as “the combination of component structures (effected by correspondences between their subparts) to form a composite structure”.

<sup>56</sup> Langlotz’s (2006) approach to variation analysis will be reviewed in more detail in the course of Chapter 4, Section 4.2.2.

<sup>57</sup> Langlotz (2006) uses the term analyzable in the same sense as Langacker (1987: 486), who defines analyzability as the extent to which the contribution of component structures to a composite structure is recognized.

- (23) **Renewable cat out of bag** *As foreseen in its 2006 energy review, the DTI has started a consultation on reform to the arrangements to support **renewable electricity** generation in the UK.*<sup>58</sup>
- (24) **Brown lets the nuclear cat out of the bag** *"We have made the decision to continue with **nuclear power**." With those ten words, Gordon Brown managed to break the law, sabotage an ongoing public consultation and do a U-turn on his promise to listen to the people - all during his first Prime Minister's Question Time.*<sup>59</sup>

Following Langlotz (2006: 269), one must expect that in the examples presented above the topic-indicators *renewable* and *nuclear* will restrict the referential scope of their head-noun *cat*. However, these adjectival premodifiers do not really define the subclass of 'information' as such, but rather modify some essential part of the TOPIC, i.e. what the information is about, e.g. 'information about the arrangements to support renewable electricity generation in the UK' in (23) above and 'information about the decision to continue with nuclear power' in (24) above. At the same time, the noun *cat* is not co-referent with head-nouns *electricity* and *power* modified by *renewable* and *nuclear* in the respective context, since it would imply paraphrases like *Renewable electricity out of concealment* in (23) and *Brown lets the nuclear power out of concealment* in (24), which is obviously not the case. Langacker's representation of the PU's semantic elements is oversimplified and does not include the element TOPIC.

(25) and (26) below are examples of what Langlotz (2006: 272) calls a topic-indicating lexical substitution when the context-specific target of a component (the topic-indicator) substitutes this component. In (25) *cat* is substituted by *SoftGnome* and in (26) by *the true feelings of the Professional Basketball Club ownership group*.

- (25) **SoftGnome is out of the bag**. *Our latest PhoneGnome add-on, SoftGnome, is now available for testing by PhoneGnome owners.*<sup>60</sup>
- (26) *Well today Chesapeake Energy Corp. Chair & CEO and Sonics investor Aubrey McClendon let the true feelings of the Professional Basketball Club*

<sup>58</sup> [http://www.reckon.co.uk/open/Renewable\\_cat\\_out\\_of\\_bag\\_-\\_viewpoint:\\_Franck](http://www.reckon.co.uk/open/Renewable_cat_out_of_bag_-_viewpoint:_Franck), Oct 11 2006

<sup>59</sup> <http://www.greenpeace.org.uk/blog/nuclear/brown-lets-the-nuclear-cat-out-of-the-bag-20070706>, Jul 6 2007

<sup>60</sup> <http://www.phonegnome.com/blog/2005/11/01/softgnome-is-out-of-the-bag>, Nov 1 2005

*ownership group out of the bag* in an interview he did with the *Oklahoma Journal Record*.<sup>61</sup>

However, substitutes of a component are not necessarily always its direct referents, cf. (27) – (30) below:

- (27) *Heroes Unmasked: How Secret Are Secret Identities? The Bat's Out of the Bag* Comic book icon: Batman<sup>62</sup>
- (28) **Potter out of the Bag:** 1,200 Readers Know Harry's Fate. A small group of lucky fans got their copy of "Harry Potter and the Deathly Hallows" early. (...) Fearing public leaks of the book's contents, Scholastic is urging fans to keep "the packages hidden" until midnight July 21, the official release date of the book, so as not to read the book early and possibly spoil the story for other readers.<sup>63</sup>
- (29) **Alfred lets 'Batman' villains out of the bag.** According to MTV's *Splash Page*, Caine, who plays Alfred the butler in the "Batman" films, told interviewers on Monday that Johnny Depp and Philip Seymour Hoffman are the choices in mind for the characters of Riddler and Penguin, respectively.<sup>64</sup>
- (30) **Return never lets Gellar out of the bag.** Rather than snow us with all manner of familiar tricks designed to help us identify with the protagonist as soon as cinematically possible, *The Return* gambles absolutely everything on intrigue. Ambiguity rules over nearly every scene and Joanna (Sarah Michelle Gellar), our heroine, is consistently held at arm's length.<sup>65</sup>

In (27) *Bat* refers to Batman's secret identity, in (28) *Potter* does not refer to Harry Potter but to his final fate revealed in the last book of the series, and '*Batman*' *villains* in (29) refer to actors who may be chosen to play villain characters. In (30) things are even more complicated: there is a reference transfer in using the name of the actress playing the role (*Gellar*) to denote the character she is playing (*Joanna*); on the other hand, it is not Joanna herself who is never let out of concealment, but her true identity.

If one compares Langacker's (1987) representation of a symbolic structure in Figure 4 to his analysis of the PU in Figure 5, one may notice that in the former the sanctioning structure corresponds to the conventional unit and the target structure to the usage event sanctioned by this unit, while in the latter both the sanctioning and the target structures have the unit status. This means that Langacker's analysis of

---

<sup>61</sup> <http://www.sportsbusinessradio.com/node/1268>, Oct 12 2009

<sup>62</sup> <http://entertainment.howstuffworks.com/secret-identity4.htm>, Mar 23 2005

<sup>63</sup> <http://abcnews.go.com/Business/FunMoney/story?id=3390560&page=1>, Jul 18 2007

<sup>64</sup> <http://www.collegiatetimes.com/stories/11830/alfred-lets-batman-villains-out-of-the-bag/print>, Sept 10 2008

<sup>65</sup> <http://www.vueweekly.com/article.php?id=5015>, Nov 14 2006



PU structure actually lacks the representation of the usage event and therefore fails to describe contextual meaning. Thus, the semantic analysis of the PU *THE CAT ... OUT OF THE BAG* presented by Langacker (1987) in Figure 5 cannot account even for simple cases of contextual reference, let alone the reference transfer examples shown in (27) – (30) above.

Langlotz (2006) attempts to improve Langacker's (1987) model by adding a usage event as a Target structure (Figure 6 below), while in the formal description of the Standard/Sanctioning structure he adapts Geeraerts' (1995) prismatic model of semantic relations in idioms (this model is discussed in more detail in Section 2.3.2 of this chapter).

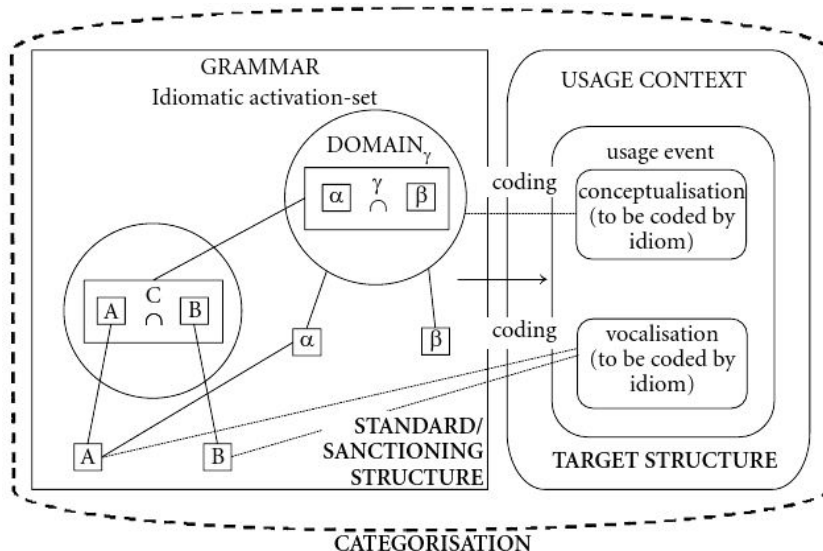


Figure 6 Idiomatic usage event (Langlotz 2006: 186)

According to the interpretation of metalanguage which Langlotz (2006: 109) himself provides, the letters A, B, C and  $\alpha$ ,  $\beta$ ,  $\gamma$  stand for semantic substructures, and the connection lines which are drawn between them indicate semantic processes. Keeping this in mind, it is unclear how Langlotz manages to draw coding links between some semantic units in the Sanctioning structure and vocalization in the Target structure, which obviously belong to different levels of representation. Even in Langacker's (1987) description (Figure 4 and Figure 5) coding takes place between the elements of the same level.

This is the most obvious inconsistency of the picture in Figure 6, but there are other problematic aspects characteristic of this model in general.

As has been pointed out by Culicover & Jackendoff (2005: 19, 531) and Jackendoff (2002b: 427), Cognitive Grammar downplays the importance of syntax in favour of semantics. The syntactic formation rules are minimized, syntax is mostly derivative from semantics and little is said about phonology at all. Jackendoff (2002) finds such rejection of syntax an unnecessary overreaction. Langlotz (2006: 75) remarks that he will only focus on the semantic pole since phonological structuring is not of direct relevance for the analysis of idioms. Such a position would make it impossible to account e.g. for substitution of the default lexical item *cat* by its homonyms and near-homonyms in (31) – (38) below:

- (31) *Where the qat<sup>66</sup> is out of the bag. It's a drug that induces dreaminess, lucidity and, later on, surges of energy. And in some countries, including Yemen, it's legal, says Brian Whitaker.*<sup>67</sup>
- (32) *Khat out of the bag. A Somali national residing in London was caught with 10 kilogrammes of khat at the Malta International Airport (MIA) last week.*<sup>68</sup>
- (33) *"The CATS<sup>69</sup> Are Out of the Bag"*<sup>70</sup>
- (34) *Keeping the CAT(echolamine)<sup>71</sup> Out of the Bag*<sup>72</sup>
- (35) *Letting the CAThellicidin<sup>73</sup> out of the bag, as a therapeutic modulator of the adaptive immune system.*<sup>74</sup>
- (36) COMEDY CENTRAL(R) *Lets the Katt Out of the Bag: The World Television Premiere of 'Katt Williams: American Hustle' Debuts Sunday, January 13 at 10:00 p.m.*<sup>75</sup>
- (37) *Web leak of Linux lets Hat out of the bag. Parts of the newest version of Red Hat's Linux software slipped onto the Internet Wednesday, nearly a week before the*

---

<sup>66</sup> *Khat, qat* ['ka:t] is a flowering plant native to tropical East Africa and the Arabian Peninsula, classified as a drug. It is a controlled/illegal substance in many countries.

<sup>67</sup> <http://www.guardian.co.uk/world/2001/may/28/worlddispatch.drugsandalcohol>, May 28 2001

<sup>68</sup> <http://www.maltatoday.com.mt/2008/05/14/n7.html>, May 14 2008

<sup>69</sup> CATS refer to catastrophe bonds (also known as cat bonds) – risk-linked securities that transfer a specified set of risks from a sponsor to investors.

<sup>70</sup> <http://www.capitallinkforum.com/cef/2008/pres/temple.pdf>, Apr 16 2008

<sup>71</sup> *Catecholamines* are chemical compounds, hormones that are released by the adrenal glands in situations of stress.

<sup>72</sup> <http://neurology.jwatch.org/cgi/content/citation/2008/708/2>, May 8 2008

<sup>73</sup> *Cathelicidin* is an antimicrobial protein found in specific granules of polymorphonuclear leukocytes.

<sup>74</sup> <http://linkinghub.elsevier.com/retrieve/pii/S0145212604003832>, Jan 23 2005

<sup>75</sup> [www.reuters.com/article/pressRelease/idUS167657+07-Jan-2008+PRN20080107](http://www.reuters.com/article/pressRelease/idUS167657+07-Jan-2008+PRN20080107), Jan 7 2008

*operating system's official release date, giving glimpses of a product with a new focus on mainstream computer users.<sup>76</sup>*

(38) *Letting<sup>[sic]</sup> the scat out of the bag: Borat's boys speak<sup>77</sup>*

### 2.3.2 Isomorphism and motivation as different dimensions of PUs' compositionality

Compositionality and non-compositionality are frequently occurring terms when it comes to PU semantics. As Nenonen (2002: 17) points out, the notion of compositionality is one of the most obscure concepts in linguistics. Generally, the concept is defined as the derivability on the meaning of the whole from the meanings of its constituents. Geeraerts (1995), who proposes to treat compositionality of idioms as various combinations of isomorphism and motivation, gives a schematic representation of semantic relations in idioms as shown in Figure 7 below.

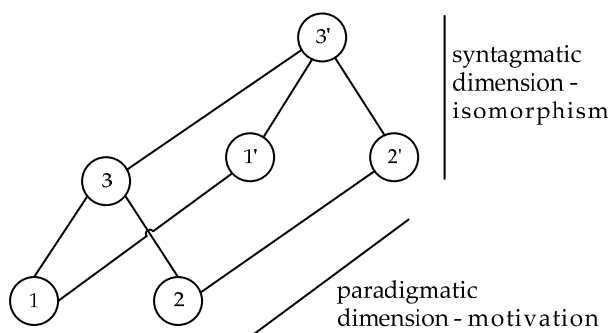


Figure 7 *Semantic relations in idioms according to Geeraerts (1995: 60)*

In Figure 7, numbers 1, 2 and 3 represent first and second constituent lexical units and expression as a whole in their literal reading, and numbers 1', 2' and 3' those in their idiomatic reading, respectively. According to Geeraerts (1995: 60-62), motivation refers to paradigmatic lines of the prism and isomorphism to syntagmatic ones, i.e. they belong to different planes. The question that remains concerns the nature of these lines. In order to answer this question one has to look at definitions of isomorphism and motivation, which Geeraerts (1995) provides.

<sup>76</sup> <http://www.zdnet.com.au/web-leak-of-linux-lets-hat-out-of-the-bag-120268556.htm>, Sep 26 2002

<sup>77</sup> <http://moviecitynews.com/2007/01/letting-the-scat-out-of-the-bag-borats-boys-speak/>, Jan 8 2001

Geeraerts (1995) defines motivation as “the transparency of the semantic extension that leads from the original meaning of an expression to its transferred reading”. Geeraerts (ibid. 61) also claims that both isomorphism and motivation involve transparency: isomorphism coincides with syntagmatic transparency and motivation with paradigmatic transparency. Based on this definition, the paradigmatic lines should represent the transition or extension from the literal to idiomatic meaning and syntagmatic lines the transition or extension from the constituent meanings to the meaning of an expression. These lines can in their turn be transparent or opaque. This definition has several problems. First of all, it is not clear what such transition or extension from one meaning to another could mean in practice and the same holds for transparency and opaqueness. As Geeraerts (ibid.) puts it, the transition from one meaning in the prismatic structure to another is opaque if the latter “cannot be derived on the basis of the former”. In this connection he also speaks of syntagmatic and paradigmatic nonderivability of meaning. However, the term derivation in relation to meaning does not add any clarity to the picture. Geeraerts (1995) claims that in the course of a syntagmatic derivational process “the meaning of a compound expression is computed on the basis of the meanings of the constituent parts of the expression”. This implies that in an idiom which is both isomorphic and motivated the idiomatic meaning of a unit as a whole can be syntagmatically computed from the paradigmatically derived meanings of its lexical components and at the same time it can be paradigmatically derived from the literal meaning of the expression, which in turn is syntagmatically computed from the literal meanings of the components. This bottom-up scenario is shown in Figure 8 below with arrows indicating the direction of derivation.

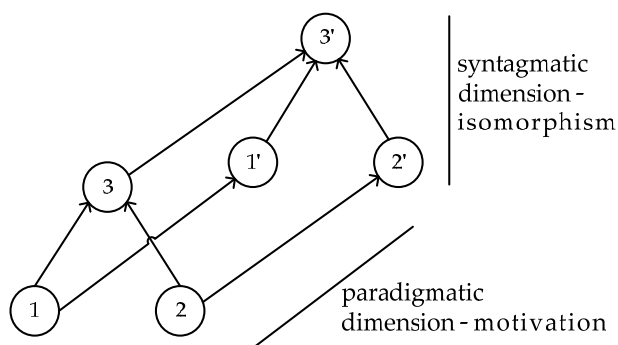


Figure 8 The bottom-up scenario Geeraerts (1995)

The only way such a bottom-up compositional process could work in practice is through the availability of independently existing paradigmatic motivation lines, i.e. if lexical components possess metaphorical meanings which are independent of the expression itself. Otherwise the idiomatic meaning could never be “computed”. In the case of the Finnish PU X HEITTÄÄ HELMIÄ SIOILLE | NP<sub>SUBJ</sub>[N{X}] V{heittää ‘throw’} NP<sub>OBJ</sub>[N<sub>PL</sub>{helmi ‘pearl’}] PP<sub>ALL</sub>[N<sub>PL</sub>{sika ‘pig’}] | lit. ‘X throws pearls to pigs’, id. ‘X causes a transfer of some entity, evaluated by the speaker as good, to a recipient, evaluated as inadequate in some way’, the input for the composition of idiomatic meaning like ‘to give something valuable to somebody who cannot appreciate it’ has to contain independent metaphorical meanings of *heittää* ‘throw’ as ‘give’, *helmi* ‘pearl’ as ‘something valuable’ and *sika* ‘pig’ as ‘somebody who cannot appreciate something valuable’. According to NS (Sadeniemi 1970), such meaning can be detected only for one of these three components – *helmi* ‘pearl’:

*Kuv. jstak arvokkaasta, korvaamattomasta, kauniista yksilöstä, esineestä tms.*

‘Fig. about some valuable, irreplaceable, beautiful person, object etc.’

Cf. examples like (39) – (41) below:

- (39) *Venäläinen kirjallisuus on tuottanut useita maailmankirjallisuuden helmiä.*

‘The Russian literature has produced several **pearls** of world literature.’

- (40) *Capri on Välimeren helmi.*

‘Capri is the **pearl** of the Mediterranean.’

- (41) *Hän on todellinen helmi sihteeriksi.*

‘He is a true **pearl** of a secretary.’

The verb *heittää* ‘to throw’ does not have a metaphorical meaning ‘to give’. *Sika* ‘pig’ has metaphorical meanings ‘somebody untidy, dirty’ and ‘somebody coarse, indecent’, but these do not correspond to the figurative reading of *pig* within the expression in question. On the other hand, *sika* ‘pig’ appears as a component in a number of Finnish proverbs and sayings where it displays different connotations (semantic associations) which this animal has received in Finnish culture, e.g. ‘incapable of doing anything properly’ in (42) and (43) or ‘non-civilized, primitive, unsophisticated, ignorant’ in (44) – (46) below (Laukkanen et al. 1978):

- (42) *Pane sika matkaan ja mene itse perässä.*

lit. ‘Send a **pig** on its way and go after it yourself.’

- (43) *Kerran sika metsään, silloinkin jalka poikki.*

lit. 'Once a **pig** (went) to the forest, one more time broke its leg.'

- (44) *Mitä sika hopealusikan päälle ymmärtää, ottaa suuhunsa ja puree rikki.*

lit. 'What does a **pig** understand about a silver spoon, takes into its mouth and bites to pieces.'

- (45) *Mitä siat kirkossa tekevät, ei siellä rankkia ole.*

lit. 'What are **pigs** doing in the church, there is no draff there.'

- (46) *Mistäs sika unta näkee, jollei kaukalostansa.*

lit. 'What can a **pig** dream of other than of its trough.'

Could these connotations be regarded as independent metaphorical meanings of the word *sika* 'pig' in modern Finnish? They could if this word could be separately used to denote someone unsophisticated or incapable of doing something. However, this is not the case. Of these two connotations 'ignorant, unsophisticated' would fit better to be mapped onto the idiomatic meaning of the Finnish PU *X HEITTÄÄ HELMIÄ SIOILLE* | *NP<sub>SUBJ</sub>[N{X}] V{heittää 'throw'} NP<sub>OBJ</sub>[N<sub>PL</sub>{helmi 'pearl'}] PP<sub>ALL</sub>[N<sub>PL</sub>{sika 'pig'}]* | lit. 'X throws pearls to pigs', id. 'X causes a transfer of some entity, evaluated by the speaker as good, to a recipient, evaluated as inadequate in some way', but the notion of semantic extension is inapplicable here, since in this case the connotation does not extend to an independent meaning, but rather has to be considered as a part of the word's semantic description<sup>78</sup>.

Geeraerts (1995: 62) himself analyses a Dutch PU *DE KOE BIJ HORENS VATTEN* lit. 'to take the cow by the horns', id. 'to tackle a problem by its most difficult aspect'. He starts from the assumption that this PU is both isomorphic and motivated. It is isomorphic because "a consistent one-to-one mapping can be defined between the elements of the global meaning and the meanings of the constituent parts of the expression". The cow maps onto the problem and the horns to the most difficult part of it. It is motivated, since "it is easy to see that the literal situation described by *DE KOE BIJ HORENS VATTEN* is a metaphorical image for tackling a problem at its most difficult spot" (Geeraerts 1995: 63). As far as motivation is concerned, I find it quite difficult to see how a metaphorical image for 'tackling a problem at its most difficult spot' could be independently derived from the literal situation of 'taking the cow by the horns' without

<sup>78</sup> Figure 47 in Chapter 3, Section 3.4.1 presents the semantic description of the lexical item *sika* 'pig'. Institutionalized attitudes and values are grouped there under the category Evaluation.

any prior knowledge of the PU's idiomatic meaning. As Keysar & Bly (1995) point out in their psycholinguistic study, our prior knowledge of the idiom's meaning affects the way we attempt to motivate it, i.e. motivation is rather a top-down post hoc inference process, which people make about the already available meaning. For instance, if we were told that the meaning of *DE KOE BIJ HORENS VATTEN* was 'to be engaged in a dangerous activity', the same literal situation would produce a different "metaphorical image".

Geeraerts (1995: 63) is also soon forced to acknowledge that this PU's idiomatic meaning cannot be arrived at via a bottom-up process illustrated in Figure 8 above, since the interpretation of *koe* 'cow' as 'problem' and *horens* 'horns' as 'most difficult aspects' cannot be reached independently, but rather:

[...] the input for a compositional process can be retrieved only when the output of the process (the global figurative meaning of the idiom as a whole) is already available (Geeraerts 1995: 63).

This statement is rather paradoxical in itself, since it implies that the output of a top-down analytical re-interpretation of the expression's components on the basis of its global idiomatic meaning is then re-used to provide an input for a bottom-up derivation of the same meaning, i.e. the process is in fact bidirectional. Geeraerts (ibid. 61) attempts to tackle this problem by suggesting a treatment of isomorphism as a nondirectional concept. He defines isomorphism as:

[...] a one-to-one correspondence between the formal structure of the expression and the structure of its semantic interpretation, in the sense that there exists systematic correlation between the parts of the semantic value of the expression as a whole and the constituent parts of that expression (Geeraerts 1995: 61).

In Section 3.4.3.3 of the following Chapter 3 this definition will be criticized from the positions of Conceptual Semantics.

All in all, non-directional correspondence links seem to be a more plausible way to represent semantic connections. As Geeraerts (ibid. 60) himself admits, his prismatic structure (Figure 8) is not a valid formal (i.e. formalized) representation of an idiom's semantic structure. On the other hand, when he talks about correspondences between the formal structure of the expression and the structure of its semantic interpretation, he uses the term formal as referring to form, i.e. syntax and phonology – something which his picture lacks. Since the prism only shows

connections between the literal and the idiomatic readings of the expression as a whole and the respective readings of its constituents, it is not clear how it can serve as a representation of isomorphism, which must also involve formal (i.e. syntactic, phonological) structure of the expression.

### 2.3.3 Motivation by conceptual metaphors

The cognitive theory of metaphor developed by Lakoff (1987) and Lakoff & Johnson (1980) offers the following notion of idiom motivation and explanation of motivational phenomena:

The relationship between A and B is *motivated* just in case there is an independently existing link L, such that A-L-B “fit together.” L *makes sense* of the relationship between A and B (Lakoff 1987: 448).

A motivating link (L) that relates an idiom (A) to its meaning (B) is claimed to be constituted by three elements (Lakoff 1987: 449, 451):

- conventional image
- knowledge associated with the image
- metaphors

These elements of the motivating link are assumed to be independent of the expression:

What it means for an idiom to “be natural” or to “make sense” is that there are independently existing elements of the conceptual system that link the idiom to its meaning (Lakoff 1987: 449).

Such understanding of motivation suffers from several problems. Firstly, there is no convincing non-linguistic evidence that metaphors exist independently of their linguistic manifestations. Lakoff’s (1987) claim that conceptual metaphors are independently existing conceptual structures is based solely on intuitions about how certain idiomatic expressions thematically cohere. It is a clear case of circular reasoning, where the linguistic evidence constitutes both the motivation for the hypothesis that metaphors transcend their linguistic manifestations and its only source of support. In order to prove the postulated independent nature of conceptual mappings one has to present evidence that is independent of the linguistic evidence (McGlone 2001: 95). As Keysar & Bly (1999: 1564) point out, Lakoff’s theory does not include criteria for negative evidence and therefore it is not falsifiable.



Secondly, there is no non-linguistic evidence that conventional images exist independently of their linguistic manifestations. Lakoff (1987: 449) claims that the literal meaning of the idiom fits the conventional image. However, it could just as well be vice versa. Like in the case of conceptual metaphors, there is no way to prove that the image associated with the idiom exists independently of the idiom's literal and idiomatic meaning. Thus, speaker-to-speaker differences in the images associated with idioms may reflect their different understanding of the idiom's meaning and/or differences in knowledge, whereas the uniformity of image among speakers is the result of the shared meaning.

According to Lakoff (1987: 449), metaphors map the literal meaning, the image and its associated knowledge into the idiomatic meaning. Nevertheless, Lakoff does not claim that the motivating link makes the idiomatic meaning predictable. He starts from the assumption that this meaning is not completely arbitrary, but at the same time he takes it for granted. However, he does not consider the possibility that the same idiom could in principle mean something different and still make sense. For instance, if we have learned that *TO KEEP SOMEONE AT ARM'S LENGTH* means 'to keep someone from becoming intimate' we can, as Lakoff (1987: 447-449) suggests, associate the image with defence, but as Keysar & Bly (1999) argue this idiom could also mean 'to be very close to a person', and in such a case the image and the knowledge associated with it would be different: the purpose of having one's arm stretched towards another person would be keeping him close and preventing him from going away. Thus, the same distance could be interpreted both as a long one and a short one, depending on which idiomatic meaning one is attempting to motivate. This demonstrates that images are the result of knowing the meaning of the idiom.

Keysar & Bly (1999: 1564f) argue that since idiomatic expressions cannot provide negative instances for a particular mapping, they cannot provide positive evidence either. If one considers all links between what an idiom could possibly mean and still make sense as positive evidence, the same expression could in principle serve as evidence for several different conceptual mappings. On the other hand, if a meaning does not make sense, it cannot be regarded as evidence that no motivating mappings exist. According to Keysar & Bly (1995, 1999), the same meaning can make sense if it is considered an official meaning of an idiom and seem opaque otherwise. Once our interpretative system

establishes a motivating link between an idiom and its meaning, other potential meanings become less sensible. While the conceptual metaphor view could account for cases where the meaning seems to be motivated, it is not able to explain why the same meaning does not make sense if it is not perceived as the stipulated meaning.

If the metaphorical link was truly independent of the expression, it would motivate the meaning regardless of whether it was stipulated or not. It would also motivate regardless of the speaker's ability to relate the literal meaning to a particular context (e.g. the hypothetical conceptual metaphor *PURPOSEFUL ACTIVITY IS BOXING* would motivate the Finnish idiom *HEITTÄÄ PYYHE KEHÄÄN* 'to throw in the towel') or activate a specific encyclopedic knowledge. If the conceptual metaphors are neither sufficient nor necessary to determine the full motivation their role as motivating links can be questioned. As a rule, there is no principled way to distinguish between cases where the conceptual metaphors indeed provide a source of motivation and those where a researcher who is predisposed to look for particular structures imposes them on idioms ad hoc.

The following points of criticism that could be raised against Lakoff's (1987) and Lakoff & Johnson's (1980) theory refer to its general assumptions and thus reflect on its claimed status of a coherent account of conceptual representation and figurative competence<sup>79</sup>.

Embodiment Hypothesis (i.e. the claim that the human conceptual system, including spatial concepts, is based on preconceptual bodily experiences and comprehended through the body) is not substantiated in a stringent manner. It received no supporting evidence from the source where it was expected to be found in the first place: as Nikanne (2004: 101-110) shows, written instructions for running technique describe this basic physical activity in spatial terms, while the vocabulary referring to the bodily experience is very small.

The theory of conceptual metaphor blurs the distinction between literal and figurative language by understanding metaphorical expressions in terms of their constituents' literal category membership and at the same time assuming that our knowledge of literal categories is metaphorical at some deep level (McGlone 2001: 107).

---

<sup>79</sup> For more criticism of Lakoff's theory see e.g. Dobrovol'skij & Piirainen (2005), Murphy (1996, 1997), Keysar & Bly (1995, 1999), and McGlone (2001).

Apart from a very vague idea of cross-domain mapping, the theory does not present any valid or explicit model of how metaphorical representations are constructed and constrained. Lakoff admits that metaphorical mappings are partial, but there is no explanation of why certain concepts are mapped from the source domain while others are not (Nikanne 1992: 68). The Invariance Principle (Lakoff 1993: 251f) implies that the target domain has an inherent structure of its own, but we still do not know what guarantees the supposed restructuring of one meaning domain in terms of the other. Any two concepts can be matched if the level of categorization is abstract enough. We have to recognize the hidden criterion on which the comparison is based (Sovran 1993: 26f).

The set of conceptual metaphors is not as “well-defined” (Langlotz 2006: 47) as it is asserted. For instance, Jackendoff & Aaron (1991: 324) wonder how it is possible to determine the appropriate level of abstraction of the mapping (e.g. LIFE IS A FIRE instead of a more specific schema LIFE IS A FLAME or a more general one LIFE IS SOMETHING THAT GIVES OFF HEAT), while Ortony (1988: 99) points out that the set of conceptual mappings is in principle not restricted and there is no plausible way to determine its completeness.

Lakoff & Johnson’s (1980) view of metaphors as pre-stored idealized cognitive models and mappings as a fixed part of our conceptual system ignores a long-established fact that the semantic interpretation process of metaphor is crucially and systematically context-dependent (Leezenberg 2001). Langlotz’s claim that conceptual metaphor models prevent nonsensical sense attributions (Langlotz 2006: 49) contradicts with an observation made by Sovran (1993: 44f) that even a novel metaphor that sounds nonsensical in isolation could become meaningful given a certain context:

An appropriate context can extract from each word that functions as a vehicle any meaning component that may connect it to any tenor. This leads to the conclusion that there are no real borderline between metaphor and nonsense but only good and bad metaphors (Sovran 1993: 45).

The theory of conceptual metaphor will be addressed once more in Chapter 4, Section 4.2.2.3.1 in the context of variation analysis. It will be argued that this theory cannot be regarded as a solid explanatory device for systematic variation of PUs.

### 2.3.4 Cognitive modelling of motivation in Conventional Figurative Language Theory (CFLT)

The central notion of Conventional Figurative Language Theory (CFLT) developed by Dobrovol'skij & Piirainen (2005) is image component (rich image<sup>80</sup>) – a specific conceptual structure evoked by the literal sense encoded in the lexical structure of a given unit (Dobrovol'skij & Piirainen 2005: 162). According to Dobrovol'skij & Piirainen (2005: 5), this element of the content plane of a conventional figurative unit (CFU) provides relevant motivational links between the lexical structure and the actual meaning (the latter term is used in the sense of figurative meaning) of a CFU. Image component is considered to be “the most salient feature of figurative language” and a second conceptual level at which CFUs “are associated with the sense denoted by their literal form”. Thus, the content plane of a CFU not only consists of the actual meaning, but also includes traces of the literal meaning inherited by the figurative meaning (ibid. 14). The authors also believe that rich images are important for explaining most linguistically relevant features of idioms (ibid. 61): the image component is often responsible for relevant restrictions in the usage of idioms, as well as semantic and pragmatic differences between synonymous and quasi-equivalent CFUs (ibid. 64f).

The authors assume that in iconically motivated<sup>81</sup> idioms motivational links can be described either on the superordinate level of the conceptual metaphor (the abstract metaphoric model) or on the basic level<sup>82</sup> of the rich image (ibid. 162). The authors state that procedures generating actual meaning of a CFU and motivating it are not based on meaning transfer, but on the activation of relevant knowledge structures and operations on them, and that “motivation phenomena are not reducible to operations on linguistic structures such as lexical constituents in their literal and non-

---

<sup>80</sup> This notion is borrowed from Lakoff (1987: 444f). The term image is not limited to visual images. The epithet “rich” refers to the image’s relative richness in detail compared to abstract schemas.

<sup>81</sup> Iconic motivation is based on similarity between the entity denoted by the actual meaning and the entity reflected by the underlying image (Dobrovol'skij & Piirainen 2005: 90).

<sup>82</sup> The notions of superordinate level and basic level are used in the sense of Rosch (1975). E.g. a *vehicle* is a superordinate level category that includes *car* as a basic level category. Lakoff (1993: 212) assumes that metaphorical mappings in general are at the superordinate level rather than the basic level. Dobrovol'skij & Piirainen (2005), on the contrary, emphasize the importance of basic level metaphors.

literal meaning” (ibid. 170, 173). Their postulation about the relevance of the image component requires further research into the cognitive operations which are involved in activating the motivational links and representing them in formal metalanguage (ibid. 165-166). Thus, the authors see as an important task of CFLT developing metalinguistic tools capable of capturing cognitive operations that would enable them to describe cases of iconic motivation. Such a metalinguistic apparatus is based on cognitive modelling, i.e. “formalizing conceptual correspondences between the interacting knowledge structures” (ibid. 165). An earlier attempt to develop a metalanguage has been made by Baranov & Dobrovol’skij (i.a. 1990; 1996; 1999). It includes the following elements (Dobrovol’skij & Piirainen 2005: 173):

1. Knowledge (Kn): Dobrovol’skij & Piirainen (2005: 4, 6) argue that the image component often preserves specific conceptual structures that they generally refer to as knowledge structures (i.e. world or encyclopedic knowledge). In iconically motivated idioms various types of knowledge are involved in the notion of rich image<sup>83</sup>. The authors make a distinction between natural experience (knowledge that is independent of culture) and cultural knowledge (ibid. 94). Most CFUs are clearly motivated by these underlying structures of knowledge (ibid. 31). Thus, in the case of metaphoric motivation the knowledge structures of the source domain are mapped onto the knowledge structures of the target domain (ibid. 109). In principle, the term knowledge structure is used by Dobrovol’skij & Piirainen (2005) as a synonym of the term frame.
2. Frame: This notion is adapted from Frame Semantics (Fillmore et al. 2003), where it is defined as “a schematic presentation of a situation type that underlies the meaning of a word (or of the members of sets of words) along with named participant roles or aspects of the situation”. As compared to the FrameNet, the peculiarity of Dobrovol’skij & Piirainen’s (2005: 163) approach is that the apparatus of frames is applied to purely semantic and conceptual parameters rather than to cognitive aspects of the figurative structure of predicates, and thus the interface between syntax and semantics, emphasized in FrameNet, is not present in CFLT. Dobrovol’skij & Piirainen use frame as an

---

<sup>83</sup> Cf. with Lakoff’s (1987: 449, 451) division of motivating links into three distinctive elements: conventional image, knowledge associated with image and conceptual metaphors, discussed in the previous Section 2.3.3.

umbrella term for both frames proper (static knowledge structures) and scripts (dynamic knowledge structures).

### 3. Subframe

4. Slot: This notion is similar to the notion of frame elements, i.e. relevant constituent parts and participants of a given situation (Fillmore et al. 2003).

### 5. Frame name

### 6. Frame content

7. Filler – “conceptual content of the slot” (Dobrovol’skij & Piirainen 2005: 163)

The actual meaning of idioms can be derived from frames with a help of a restricted set of the following conceptual operations:

1. Introduction of a cognitive structure into another one:  $\text{Into}(\text{CS}_1, \text{CS}_2)$
2. Elimination of a cognitive structure from another one:  $\text{El}(\text{CS}_1, \text{CS}_2)$
3. Highlighting of a cognitive structure:  $\text{Hl}(\text{CS})$
4. Repetition of a cognitive structure:  $\text{Rep}(\text{CS})$

These operations are implemented on the slots of frames and their subframes (i.e. on knowledge structures) and reveal motivating links relevant to a given idiom (Dobrovol’skij & Piirainen 2005: 174). I.e. motivating links are not posited between the literal meaning and the actual meaning, but rather between the latter and “the mental image fixed in the lexical structure” of an idiom (ibid. 17).

Thus, the actual meaning of an iconically motivated CFU is interpreted in CFLT as source-target correspondence. For example, the lexical structure of the German CFU *DAS FÜNFTE RAD AM WAGEN (SEIN)* lit. ‘(to be) the fifth wheel on the coach’, id. ‘(to be) someone who is with a group of people even though that group does not want or feel comfortable with them’ activates the source frame CAR, the filler of the relevant slot (the number of wheels) of the source frame is highlighted and replaced by a non-characteristic filler (five wheels) and, finally, mapped onto the corresponding slot (the uncomfortable person) of the target frame GROUP OF PEOPLE (Baranov & Dobrovol’skij 1996).

Dobrovol’skij & Piirainen (2005) suggest that new principles of semantic explanation that would enable taking the source frame into account should be developed. Metalinguistic instruments are supposed to be used for meaning explanation with regard to the image component. However, the authors seem to make a difference between formalized and explicit semantic explanation (Dobrovol’skij & Piirainen 2005: 166f).

On the one hand, Dobrovol'skij & Piirainen (2005) acknowledge that a formalized representation of a semantic structure is a necessary step of its accurate analysis. On the other hand, their distinction between explicit and implicit ways of pointing to the image does not involve any formal ways of analysis and remains on the level of meaning paraphrasing:

- The implicit strategy requires the distribution of relevant semantic information among various elements of the meaning explanation.
- The explicit strategy includes semantic operators ('like X', 'is perceived as X', 'is associated with X', 'is analogous with X') introducing a special part of the meaning explanation (e.g. the italicized part of the definition in (47) below), which reflects the image component, thus explicating the idea of comparison of the target structure with the relevant parts of the image and providing "the relevant motivating link" between them (ibid. 166).

Dobrovol'skij & Piirainen (2005) consider the definition of the German CFU *DAS FÜNFTE RAD AM WAGEN (SEIN)* lit. '(to be) the fifth wheel on the coach' presented in (47) below to be the explicit strategy of representing the image component of this CFU, while the definition of the Russian CFU *НОСИТЬ ВОДУ ПЕШЕТОМ/В ПЕШЕТЕ* lit. 'to carry water with/in a sieve' in (48) below is supposed to be an example of the implicit strategy:

- (47) '(to be) someone who is with a group of people, even though that group does not want or feel comfortable with them *considering their presence as redundant, as an obstacle to the normal functioning of their group*'
- (48) 'to try to achieve a goal using a totally inappropriate means for achieving this goal, which inevitably leads to failure'

The functioning of the formal metalanguage is demonstrated while analyzing motivation in the English idiom *BLACK SHEEP* id. 'someone who is considered embarrassing by the other members of their family or group because he/she is less successful or less moral than the rest' (Dobrovol'skij & Piirainen 2005: 173-181). In terms of cognitive modelling, the actual meaning is a result of interaction (metaphorical mapping) between two frames: the frame of a flock of sheep (the source concept) and the frame of a group of people (the target concept), with their respective subframes: a sheep and a person (Figure 9 below).

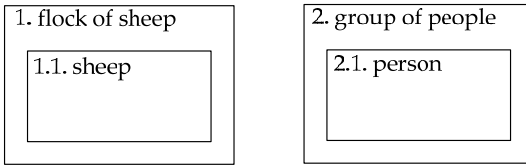


Figure 9 The source and the target concepts for the English idiom BLACK SHEEP (Dobrovol'skij & Piirainen 2005)

The first operation highlights (Hl) knowledge (Kn) in the slot *typical colour* in subframe 1.1 (Figure 10 below).

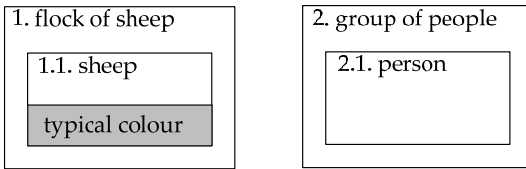


Figure 10 Operation 1:  $HlKn(Kn1[white], Slot1[typical\ colour], Subframe1.1.[sheep])$  (Dobrovol'skij & Piirainen 2005)

The second operation eliminates (El) the filler *[white]* in the slot *typical colour* (Figure 11 below).

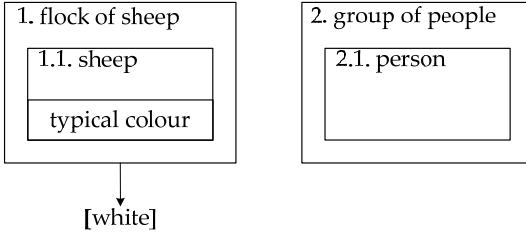


Figure 11 Operation 2:  $ElKn(Kn1[white], Slot1[typical\ colour], Subframe1.1.[sheep])$  (Dobrovol'skij & Piirainen 2005)

The third operation introduces (IntKn) the uncharacteristic filler *[black]* into the same slot of the same subframe (Figure 12 below).

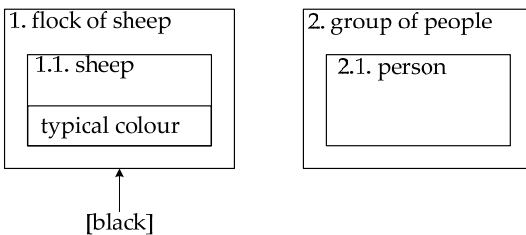


Figure 12 Operation 3:  $IntKn(Kn1[=0], Kn2[untypical\ colour:black], Slot1[typical\ colour], Subframe1.1.[sheep])$  (Dobrovol'skij & Piirainen 2005)



The fourth operation highlights the slot-content *typical colour* in the frame *flock of sheep* (Figure 13 below).

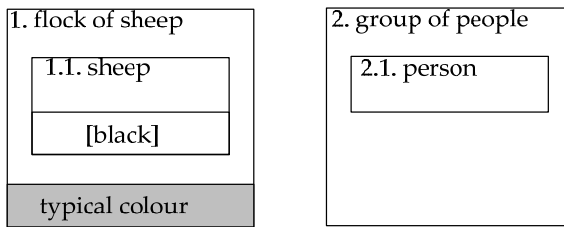


Figure 13 Operation 4:  $HLKn(Kn1[white], Slot2[typical\ colour], Frame1[flock\ sheep])$  (Dobrovol'skij & Piirainen 2005)

The fifth operation introduces the uncharacteristic filler *[black]* into the slot *typical colour* of the frame *flock of sheep* along with its typical content *[white]* (Figure 14 below).

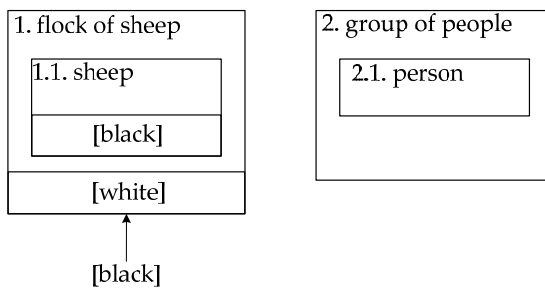


Figure 14 Operation 5:  $IntKn(Kn3[white], Kn1[untypical\ colour: black] Slot2[typical\ colour], Frame1[flock\ sheep])$  (Dobrovol'skij & Piirainen 2005)

Figure 15 below represents the result of the previous operation: a typical and an untypical colour coexist in frame 1.

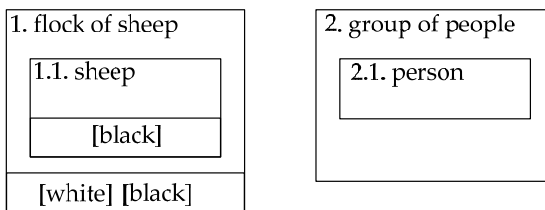


Figure 15 The result of Operation 5 presented in Figure 14 above (Dobrovol'skij & Piirainen 2005)

During the sixth operation the presence of an untypical colour in the slot *typical colour* causes the elimination of the filler *[resemblance]* from the slot *general characteristics* of the frame *flock of sheep* (Figure 16 below).

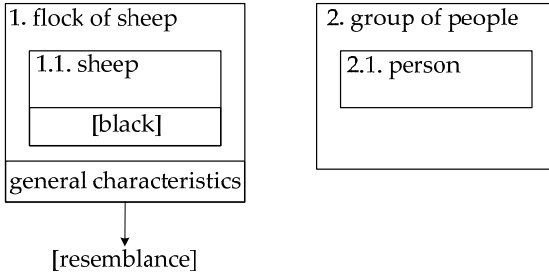


Figure 16 Operation 6:  $ElKn(Kn4[resemblance], Slot3[general characteristics], Frame1[flock sheep])$  (Dobrovol'skij & Piirainen 2005)

The seventh operation replaces resemblance as one of the *general characteristics* of a *flock of sheep* by the concept of dissimilitude (Figure 17 below).

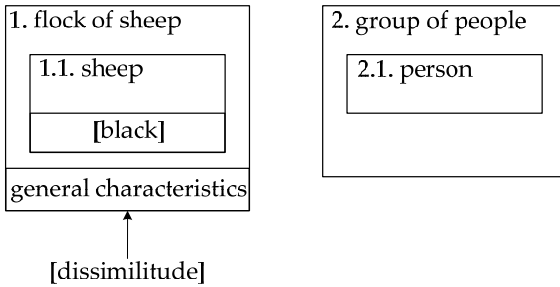


Figure 17 Operation 7:  $IntKn(Kn4[=0], Kn5[dissimilitude], Slot3[general characteristics], Frame1[flock sheep])$  (Dobrovol'skij & Piirainen 2005)

The eighth operation (Figure 18 below) highlights subframe 1.1. (*the untypical representative of a flock of sheep*).

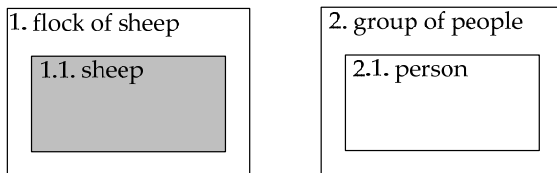


Figure 18 Operation 8:  $HlSubframe(Subframe1.1[sheep], Frame1[flock sheep])$  (Dobrovol'skij & Piirainen 2005)

After that, almost all of the above-described operations are applied in the opposite order to frame 2 *group of people*. Dobrovol'skij & Piirainen (2005) list them under the single ninth operation, which is formalized as:

RepOp(Op [Op= HlSubframe (repetition of operation 8), ElKn (repetition of operation 6), IntKn (repetition of operation 7), HlKn (repetition of operation

4), ElKn (repetition of operation 2), IntKn (repetition of operation 3)], Frame1[flock of sheep], Frame2[group of people]).

Taken separately it results in *Operations 9<sub>1</sub>-9<sub>6</sub>* (Figure 19 through Figure 24 below). During *Operation 9<sub>1</sub>* the subframe *person* is highlighted (Figure 19). *Operation 9<sub>2</sub>* eliminates the filler [resemblance] from the slot *general characteristics* in the frame group of people (Figure 20). *Operation 9<sub>3</sub>* replaces the concept of resemblance by the one of dissimilitude in the slot *general characteristics* of a *group of people* (Figure 21). *Operation 9<sub>4</sub>* highlights the filler [typical features] in the slot *general characteristics* of subframe 2.1 (Figure 22). *Operation 9<sub>5</sub>* eliminates the filler [typical features] from the slot *general qualities* in subframe 2.1 (Figure 23). *Operation 9<sub>6</sub>* introduces the filler [untypical features] into the slot *general characteristics* of subframe 2.1 (Figure 24).

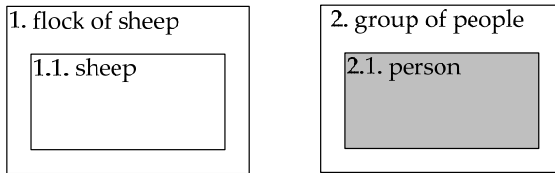


Figure 19 *Operation 9<sub>1</sub>: HlSub-Frame(Subframe2.1[person], Frame2[group of people])* (Dobrovol'skij & Piirainen 2005)

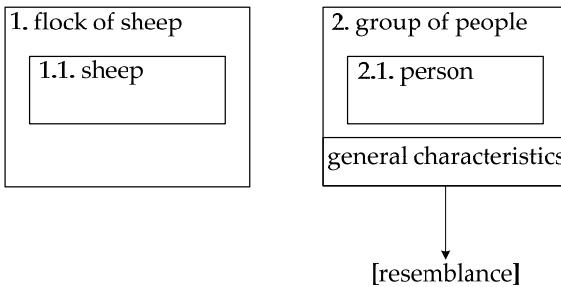


Figure 20 *Operation 9<sub>2</sub>: ElKn(Kn4[resemblance], Slot3[general characteristics], Frame2[group of people])* (Dobrovol'skij & Piirainen 2005)

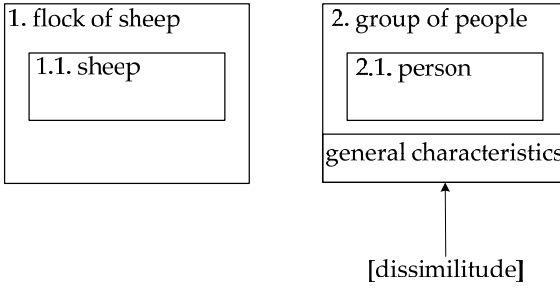


Figure 21 Operation 9<sub>3</sub>:  $IntKn(Kn4[=0], Kn5[dissimilitude], Slot3[general characteristics], Frame2[group of people])$  (Dobrovol'skij & Piirainen 2005)

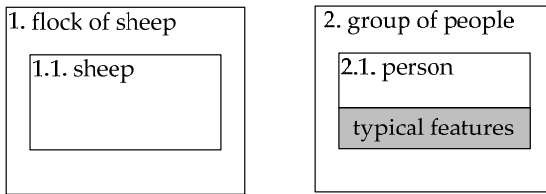


Figure 22 Operation 9<sub>4</sub>:  $HLKn(Kn6[typical features], Slot4[general characteristics], Subframe2.1[person])$  (Dobrovol'skij & Piirainen 2005)

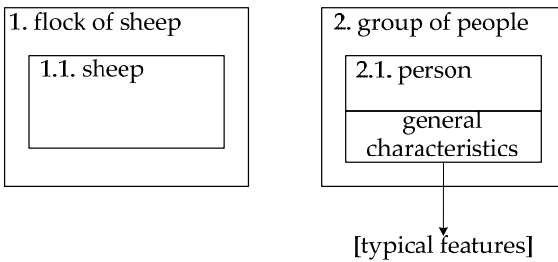


Figure 23 Operation 9<sub>5</sub>:  $ELKn(Kn6[typical features], Slot4[general characteristics], Subframe2.1[person])$  (Dobrovol'skij & Piirainen 2005)

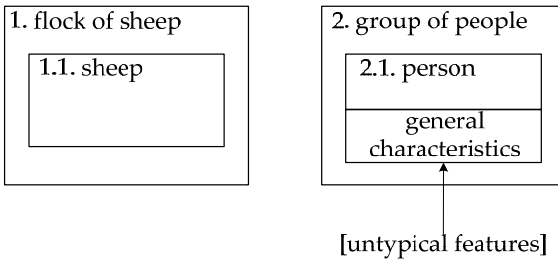


Figure 24 Operation 9<sub>6</sub>:  $IntKn(Kn6[=0], Kn7[untypical features], Slot4[general characteristics], Subframe2.1[person])$  (Dobrovol'skij & Piirainen 2005)

Finally, during the tenth operation untypical features are highlighted in the subframe *person* (Figure 25 below).

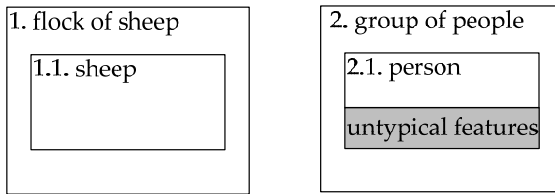


Figure 25 Operation 10: *HLKn(Kn7[untypical features], Slot4[general characteristics], Subframe2.1[person])* (Dobrovol'skij & Piirainen 2005)

Figures 9 – 25 above demonstrate the frame-based metalanguage, which Dobrovol'skij & Piirainen (2005: 164) call the cognitive modelling of the actual (figurative) meaning of idioms. According to the authors (ibid. 181), this metalanguage has a stronger explanatory power than descriptions based on meaning transfer because it models the inference process rather than points to its existence. However, it is important to mention that this model does not aim to be psychologically plausible, but is rather oriented towards explaining potential motivation links (ibid. 165). Since the authors themselves admit that motivation is a subjective criterion that cannot be verified or operationalized in a strict way (ibid. 80f), the model cannot in principle be taken to describe some objective linguistic reality.

The CFLT approach to idiom motivation is an attempt to combine Lakoff's (1987) cognitive theory of metaphor with frame-semantic analysis à la Fillmore et al. (2003). All basic notions (such as rich image and knowledge) are borrowed from Lakoff's (1987) theory. The theory itself is adapted in such a way that the superordinate level metaphoric models are assumed to determine the direction of mapping, while the rich image of the basic level metaphor is taken to provide semantic material for constructing the actual meaning<sup>84</sup> (Dobrovol'skij & Piirainen 2005: 172). This semantic material is organized and represented in the form of frames. The main fallacy here (just like in Lakoff's 1987 approach to motivation described in Section 2.3.3 of this chapter) is that the semantic material of the rich image is taken to be evoked only by the lexical structure independently of the actual meaning onto which the

<sup>84</sup> Nevertheless, Dobrovol'skij & Piirainen (2005: 172) admit that in many cases this kind of alliance is not especially productive.

image is mapped in the course of motivation. I.e., just like Lakoff (1987), the authors ignore the fact that the structure of the image is inevitably influenced by the structure of the figurative meaning.

## **2.4 PUs and context**

### **2.4.1 Psycholinguistic studies**

The prevailing tendency in psycholinguistic idiom studies is to acknowledge the importance of interaction between the PU and its linguistic context<sup>85</sup>. It is widely accepted that the result of this interaction can affect PU recognition, comprehension and interpretation (Gernsbacher and Robertson 1999, Ortony et al, 1978), especially for ambiguous PUs (Colombo 1993, Colombo 1998, Peterson et al. 2001, Cacciari et al. 2005). In this connection, the effect of context length (Schweigert and Moates 1988) and type (McGlone et.al. 1994) can also be examined. However, the majority of psycholinguistic research concerning context effects in idiom processing is aimed at showing whether a biasing context primes either the “literal” or the idiomatic interpretation of PUs by affecting the temporal course of activation of these meanings, i.e. by making one of them available more quickly (different models of context vs. meaning activation are described in Colombo 1993). As far as I know, the fact that idiomatic interpretation itself can vary depending on the context has not been the subject of any serious psycholinguistic investigation.

### **2.4.2 Sufficient context, cohesion**

Apparently, contextual interpretation of PUs is secured by their semantic relation to other elements of context. A problem of sufficient context arises in relation to this. The presence of all necessary elements in the local context, as well as the proper understanding and interpretation of the intended meaning by the addressee, is not necessarily guaranteed by the speaker, although Dobrydneva (2000: 96) includes this criterion in her definition of phraseological context. As Ortony et al. (1978: 476) remark, “there certainly are cases, where an utterance is insufficiently related to the context for it to be understood”.

---

<sup>85</sup> Although Gibbs (1980) tries to downplay the role of context in the interpretation of PUs.

A particularly interesting question is how this interpretation emerges in cases of variation (both formal and semantic). Naciscione (2001: 41, 47-52) claims that it is facilitated by phraseological cohesion, defined as:

[...] part of the meaning of the base form; the unity of phraseological meaning in instantial use; a semantic and stylistic relation, realized in discourse by virtue of ties with the base components (Naciscione 2001: 51).

Cohesion presupposes that the interpretation of a particular element of the text is dependent on that of another, i.e. one cannot be effectively decoded except by recourse to the other (Halliday & Hasan 1976: 4). Naciscione (2001: 50) points out that phraseological cohesion provides a link between phraseological components in discourse, while components themselves are dependent on the PU and the context. She concludes that the totality of these dependencies results in cohesion.

### 2.4.3 Phraseological context

Within the framework of Russian phraseological theory, the problem of PUs in relation to their context of realization has been studied by several scholars. In what follows, I will review some of them.

Dobrydneva (2000) and Avdeeva (2004) both speak of two types of phraseological context, namely the language phraseological context and the speech phraseological context. However, they define these notions differently. By the language phraseological context Dobrydneva (2000) understands:

[...] such a fragment of text in the linear space of which lexico-grammatical surrounding of the phraseological unit due to its systemically conditioned properties 'supports' and at the same time constrains the potential variety of possible realizations of phraseological meaning (Dobrydneva 2000: 96).

However, Dobrydneva (2000) does not specify what exactly these properties are and how they constrain or support meaning. Speech phraseological context, according to Dobrydneva (2000: 96), is an atypical, uncharacteristic surrounding of a PU which modifies its systemic meaning and subjects it to occasional semantic-pragmatic variation. From Avdeeva's (2004) point of view, the language context of a PU is an

abstract typical context that supports its base meaning<sup>86</sup>, while the speech context is any concrete context which surrounds a PU when it is used in speech. Avdeeva's (2004) distinction between possible and actual contexts is to a large extent based on Saussure's distinction between *langue* and *parole*, while Dobrydneva's (2000) definition of contexts is concentrated on juxtaposition of contexts with typical vs. occasional realization of PU meaning.

Avdeeva (2004) also distinguishes between strong and weak contextual positions of a PU. She calls strong such a position in which the context supports the base meaning of a PU. In a weak contextual position, the context either supports the PU's secondary meaning or does not support any of its meanings at all. Thus, the strong position is characteristic of the language phraseological context. Avdeeva (2004) divides speech contexts into several categories: usual and occasional, those that support the base meaning in the strong contextual position and those that do not support the base meaning in the weak position, and materially expressed and zero contexts. Contexts that do not support the base meaning of the PU she divides into two types: those that support the PU's secondary meaning and those that do not support any of the PU's meanings.

Avdeeva's (2004) notion of zero context can be compared to the notion of *con-situation* defined by Dobrydneva (2000: 97) as a "hypothetical, materially not expressed context", which leads to "situational conditions of actualization and accordingly interpretation of PU's semantics". As Dobrydneva (2000) states, *con-situation* is often used to identify the meaning of a PU in dialogical speech, when the turn containing a PU has no direct semantic or grammatical connection with the immediate surroundings. As it will be demonstrated in Section 5.2.2 of the present book, such delimitation of the notion of context to the immediate surroundings of one turn does not correspond to my understanding of the local linguistic context, which also encompasses turns prior to the utterance in focus if they are presented in the body of the same message as quotations. If they are not quoted, they can still be treated as a wider linguistic context, and since they are in any case materially expressed they cannot be concerned as a mere situation. This is also the reason why

---

<sup>86</sup> The terms *systemic meaning* and *base meaning* are synonymous and are used to denote the default, unmarked meaning of a PU.



Dobrydneva's (2000) distinction between situational-contextual and situational types of meaning realization cannot be applied to the purposes of my analysis. Dobrydneva (2000) assumes that in cases where the materially expressed context is absent, the speaker relies on the addressee's knowledge of the specific situation. In my case it can also mean that the relevant context probably exists as a wider linguistic context known to the members of the given community, but not recognized by an outsider.

Both Dobrydneva's (2000) and Avdeeva's (2004) definitions of phraseological context described above suffer from certain vagueness since they only postulate the supporting role of the context without actually trying to explicate the cohesive ties in the text. An attempt to determine the relationship to the PU to the context can be found in Kunin (1996: 199), who uses the notion of a phraseological actualizer – a word, phrase, sentence or group of sentences that is semantically connected with a PU in the same context. Kunin's (1996) idea is that actualizers introduce a PU into the discourse, while Dobrydneva (2000), who also uses this notion, asserts that they clarify the phraseological meaning. This approach also has several shortcomings. Firstly, there is no clear idea of what an actualizer really is and how it can be determined. Secondly, the semantic linking between the PU and its actualizer is only postulated but is not described properly. In Kunin's (1996) theory external ties of a PU in the text gain the name of its distribution, while the structural-semantic and stylistic unity formed by the PU and its actualizer is called a phraseological configuration; the latter term is also used by Dobrydneva (2000: 101) and Melerovič & Mokienko (2001: 32). However, Kunin's (1996) typology of PU's distribution only contains different syntactic configurations that a PU and its actualizer can form in the text. He distinguishes five types of distribution depending on the syntactic character of these links. Purely semantic, cohesive linking is thus not taken into account. These types are<sup>87</sup> (Kunin 1996: 200-202):

1. Compatibility which has two subtypes:
  - a. Contact compatibility – asyndetic coordination of a PU with a word or phrase which is its actualizer both in preposition and postposition, e.g.:

---

<sup>87</sup> All examples with English PUs presented in (49) – (55) are Kunin's (1996) original ones. A PU is given in bold. Actualizers are underlined only in case Kunin (1996) explicitly labels these parts of the text as actualizers.

(49) *A desire to talk **till the cows come home**.*

In the above example the verb *to talk* is an actualizer to the PU *TILL THE COWS COME HOME* 'for all time, forever'.

- b. Distant compatibility – a position in which a PU and its actualizer are separated by a word, phrase or punctuation mark, e.g.:

(50) *White-collar workers and factory hands live in adjoining blocks in houses built on the same pattern and **as like as two peas**.*

Here the PU *as like as two peas* and its actualizer *houses* are separated by the phrase *built on the same pattern*.

2. Interrelationship – a contact or distant position in which a PU is used depending on an account of a situation. This type concerns only interjectional PUs and PUs with modal meaning, e.g.:

(51) *I'll be ready very soon. – Ready **my foot**! You are still not dressed.*

3. Concatenation – a syndetic coordination of sentences, including PUs with sentence structure, e.g.:

(52) *It may be small pay, but **a bird in the hand is worth two in the bush**.*

4. Connection – subordination of sentences, including PUs with sentence structure, within the limits of a compound sentence.

(53) *It is literally true in the systematised roguery in London that **birds of feather flock together**.*

5. Asyndetic connection of two or more sentences without coordination or subordination.

(54) *Florrie: Fat chance I've got of going to France now.*

*Bessie: It's a long lane that has no turning.*

As one can see, although Kunin (1996) postulates semantic connection between a PU and its actualizer, his syntactic typology presented above is in fact not able to demonstrate any kind of semantic relation.

Kunin (1996) also distinguishes three types of phraseological context:

1. Intraphrasal – a PU and its actualizer expressed by a word or phrase within the limits of a simple or compound sentence.
2. Phrasal – a PU and its actualizer expressed by a simple or compound sentence.
3. Superphrasal – a PU and its actualizer expressed by two or more simple or compound sentences, e.g.:

(55) *Both knew there was no chance of saving the patient. They were simply **going through the motions**.*

Kunin (1996) emphasizes that superphrasal phraseological context is a complex syntactic whole. It consists of sentences that are united with respect to their semantics and syntactics. What Kunin (1996) does not explain is how these semantic links work, i.e. how a PU's referential interrelationship with the relevant parts of texts is established.

Melerovič & Mokienko (2001: 32-35) place more emphasis on the semantic aspects of phraseological configurations. For this purpose they use the term *semantization* of a PU in a text, which is defined as "revealing of the PU's actual sense by means of author's commentary". Linguistic means which are used for this purpose (words, phrases, clauses), represent a kind of phraseological actualizer, called the *semantizing phraseological actualizer* (ibid. 33). Semantization can be achieved e.g. by using periphrasis, synonymous expressions or parallel constructions. Melerovič & Mokienko (ibid. 34) remark that identification of specific elements of phraseological meaning in the context can also be made through words and phrases different from the PU in categorical-grammatical aspects but with a similar subject-logical content.

To summarize, on the one hand an explicit description of cohesive links in the text is necessary for the proper account of recognition and interpretation of a PU's contextual meaning. On the other hand, such description cannot be performed without any explicit formal description of the PU's structure, which can then be then applied to its context. Thus, one needs a model of formal semantic analysis that could explicate semantic cohesion in the text, i.e. a model that could identify and overtly relate the parts of the text that are relevant for the interpretation of the PU to parts of its semantic structure. So far, practically no attempts have been made to formalize this kind of relationship. In the course of Chapter 3, Sections 3.4.3.4 - 3.4.3.5 and Chapter 5, Section 5.4 I will present my approach to how the referential elements can be determined and formalized.

## **2.5 Negation and negative modality in PUs**

In this section I will give a brief critical overview of several previous studies of negation in Russian (Tronenko 2003), English (Palacios Martínez 1999; Moon 1998) and Finnish (Sadeniemi 1946; Kiuru 1977; Nenonen 2002) PUs. The majority of them, apart from Palacios Martínez

and Tronenko, do not attempt to provide any thorough elaboration of the phenomenon.

Tronenko (2003: 189-207) devotes a whole chapter of her book to negation in Russian phrasal idioms, which she investigates within the framework of the GB theory, trying to find systematic correspondences between negation, which is both internal (compulsory) and external (optional) to phrasal idioms, and different levels of transformational frozenness hierarchy propounded by her in the same book. However, the assignment of idioms to this or that level in this hierarchy is ruled solely by the scholar's intuitive assertions about availability or non-availability of idiomatic interpretation resulting from application of a certain modification to a certain idiom. This is also true of her statements about negation. For instance, Tronenko (ibid. 194) claims that Russian VP idioms with the negative *не* [*ne*] 'not' "exhibit consistency in terms of their object NPs being genitive case-marked". If the object NP fails to be assigned a genitive case it will block idiomatic interpretation. This claim can be proved incorrect if one turns to empirical data. The Russian phrasal PU in (56) below can be used as an example:

- (56) (С КЕМ-ТО) КАШИ НЕ СВАРИШЬ  
 [*s kem-to káši ne sváriš'*]  
 (with smb) porridge<sub>GEN</sub> NEG cook<sub>SG2</sub>  
 lit. '[one] won't cook porridge (with smb)'  
 id. 'one won't get anywhere/get along with somebody'

The search for this item in Google Groups on 14 Nov 2006 gave 213 hits with genitive negation and 25 with the accusative. Although it is obvious that genitive case is more frequent and therefore preferential for this idiom, in 24 of 25 cases of accusative case-marked NP the idiomatic interpretation was still available, e.g.:

- (57) Во-вторых, похоже с Вами кашу не сварить.<sup>88</sup>  
 [*Vo-vtorýh, pohóže s Vámi kášu ne sváriš'*]  
 lit. 'Secondly, it looks as if one won't cook porridge<sub>ACC</sub> with you.'  
 id. 'Secondly, it looks as if one won't get anywhere with you.'

Tronenko (2003) asserts that the only Russian phrasal idiom where idiomatic NP can be either marked with the genitive or accusative case is:

- (58) (КОМУ-ТО) ПАЛЬЦА/ПАЛЕЦ В РОТ НЕ КЛАД-И  
 [*komu-to pál'ca/pálec v rot ne kladí*]

---

<sup>88</sup> Google Groups: relcom.sci.philosophy/msg/c2b665241edd16d4, Aug 14 2003

(to smb) finger<sub>GEN/ACC</sub> in mouth NEG put<sub>IMPV SG2</sub>  
 lit. 'don't put a finger in (smb's) mouth'  
 id. 'smb should not be trusted'

However, Tronenko's statement (ibid. 194) that genitive marking is preferential for this item is questionable, as the search in Google Groups on 14 Nov 2006 for ПАЛЬЦА В РОТ НЕ КЛАДИ [*pál'ca v rot ne kladí*] 'finger<sub>GEN</sub> in mouth NEG put<sub>IMP SG2'</sub> and ПАЛЕЦ В РОТ НЕ КЛАДИ [*pálec v rot ne kladí*] 'finger<sub>ACC</sub> in mouth NEG put<sub>IMP SG2'</sub> gave 51 and 441 hits respectively.

Another assertion made by Tronenko (2003) is that idiomatic reading will also be eliminated by the non-occurrence of the negative *не* [*ne*] 'not' in idioms like (58) above. The author neither assumes that the idiom can occur in implicitly negative constructions like in (59) and (60) below nor takes into consideration the possibility of *не* [*ne*] 'not' to be substituted with *едва ли* [*edvá li*] 'unlikely' like in (61) below, let alone cases of creative use when the idiom in the positive form expresses a positive but nevertheless idiomatic meaning like in (62) below:

(59) В этом году даже ниче брать не надо было – у них и палатка на нас была, и все прочее, так с моими дамами разве “кашу сварить”... :((( Жена уперлась – “не хочу с котелками у костра возиться” :)<sup>89</sup>

[*V étom godú dáže ničó brat' ne nádo býlo – u nih i palátka na nas bylá, i vs'ó próčeje, tak s moími dámmami rázve “kášu svářiš”... Žená up'órlas' – “ne hočú s kotelkámí u kostrá vozít's'a”*]

lit. 'This year we didn't even have to take anything – they both had a tent for us and everything else, but will you really “cook porridge<sub>ACC</sub>” with my ladies... :((( The wife jibbed [at this] – “I don't want to mess about with kettles by the fire” :)'

id. 'This year one didn't even have to take anything – they both had a tent for us and everything else, but will you really get anywhere with my ladies...'

(60) Рената, ну со сломанным человеком – какую кашу сварить?<sup>90</sup>

[*Renáta, nu so slómanným čelověkom – kakúju kášu svářiš?*]

lit. 'Well, Renata, what kind of porridge<sub>ACC</sub> will you cook with a broken person?'

id. 'Well, Renata, how can you get along with a broken person?'

(61) Он заявил, что ракетчики его не понимают и что “с ними едва ли сварить кашу”.<sup>91</sup>

<sup>89</sup> Google Groups: relcom.wheels/msg/aae425d3b7d75ece, Sep 25 2001

<sup>90</sup> [http://gazeta.aif.ru/online/aif/1245/02\\_01?comment&Page=3](http://gazeta.aif.ru/online/aif/1245/02_01?comment&Page=3), Sep 15 2004

<sup>91</sup> <http://epizodsspace.testpilot.ru/bibl/kamanin/kniga1/02-61.html>, Sep 3 2008

[On *zajavil*, *čto rakétčiki egó ne ponimájut i čto "s ními edvá li **svářiš kášu***]

lit. 'He declared that missile specialists don't understand him and that "one is **unlikely to cook porridge**<sub>ACC</sub> with them".'

id. 'He declared that missile specialists don't understand him and that "one is unlikely to get anywhere with them".'

- (62) *Выхожу из кабинета главного редактора с мыслью, что с нашим новым начальством, пожалуй, **кашу сварю**...*<sup>92</sup>

[*Vyhožú iz kabinéta glávnoho redáktora s mýsl'ju, čto s nášim nóvym načál'stvom, požáluj, **kášu sváriš**...*]

lit. 'I step outside the editor-in-chief's office with the thought that, perhaps, one **will cook porridge**<sub>ACC</sub> with our new boss.'

id. 'I step outside the editor-in-chief's office with the thought that, perhaps, it is possible to get along with our new boss.'

In his article Palacios Martínez (1999) examines English negative polarity idioms (NPIDs). The author's objective is "to concentrate specifically on idioms with negative polarity and to examine them in close detail from both a semantic and a syntactic perspective" (Palacios Martínez 1999: 66). Palacios Martínez (1999) classifies NPIDs according to the type of negation (clause or constituent/subclause), syntactic patterns (7 different patterns for clause negation type), and semantic features (26 main categories, like *mind*, *body*, *places*, *animals* etc., subdivided into 125 subclasses). According to his definition NPIDs are "idiomatic constructions which, because of their nature, always occur in the negative form and express a negative meaning" (ibid. 65). Unfortunately, not all of the idiomatic expressions which Palacios Martínez (1999) chooses to call NPIDs, fulfil the criteria set by his definition. To give just a couple of examples: in the idiomatic meaning of *ROME WAS NOT BUILT IN A DAY* 'time, patience and hard work are needed for a difficult, or important, undertaking' (LDCE: 483) there is no indication of either explicit or implicit negation, while *DON'T PUT ALL YOUR EGGS IN ONE BASKET* is listed in ODEI as *PUT ALL (ONE'S) EGGS IN ONE BASKET* and its meaning 'to make all plans depend on the success of one thing' (Cowie et al. 1993: 351) does not contain any negative element either. Furthermore, even if one found an idiomatic construction listed in the dictionary in the negative form and expressing a negative meaning, one would have to prove that it always occurred in the negative form, or at least that its positive counterparts were cases of a very rare creative use. For that one would need empirical

<sup>92</sup> <http://www.hip-hop.ru/forum/post1282246-n321/>, Dec 21 2004

evidence, not intuition. Such evidence can be obtained only from large corpora.

Moon (1998: 106–107), in her corpus-based account of English fixed expressions and idioms (FEIs), devotes a couple of pages to polarity. Conventionally negative FEIs according to Moon (*ibid.* 106) have negative as part of the canonical expression. Unlike Palacios Martínez (1999), Moon (1998) is careful enough not to claim that these expressions will always occur in the negative form. Such FEIs make up around 5% of a total of 6776 FEIs in her database. Similar figures were given for Spanish and English idioms by Palacios Martínez (1999) and for Finnish by Nenonen (2002). Moon (1998: 106) notes that there are also some FEIs that “typically or mandatorily occur in (broad) negative environments” and thus are “more commonly negative than positive”, but gives no statistical evidence to support this claim. Moon (*ibid.* 106) also makes an interesting observation: in cases when proverbs with a negative imperative or modal in their canonical form are transformed to positive predicates, a negative evaluation may still be implied. However it is not clear what the canonical form is and what the transformation for FEIs like *HAVE ONE’S CAKE AND EAT IT*, *MAKE A SILK PURSE OUT OF A SOW’S EAR* and *PUT THE CART BEFORE THE HORSE* is. While Moon (1998) uses these FEIs as examples of transformations from negative *DON’T...*, *YOU CAN’T...*, several dictionaries will list the positive variants as base forms (Cowie et al. 1983, Kunin 1956, LDCE).

In another chapter, which is devoted to the FEI’s abilities to collocate with certain categories, structures, or lexical items, Moon (1998: 116–119) remarks that in cases where an FEI typically co-occurs with modals it is difficult to separate its meaning from the modality of the co-text (*ibid.* 117). Moon’s (*ibid.* 117) case study of the FEI *ROCK THE BOAT* showed a very strong pattern of it “being used in negative contexts, typically with expressions of the improbability, inadvisability, or undesirability”, which is a very interesting observation by itself, but unfortunately Moon (1998) does not attempt to come to any conclusion on whether this pattern expresses modality which is a part of the FEI’s meaning, or if it only comes from the co-text.

Sadeniemi (1946), in his article devoted to negative and positive polarity in the Finnish language in general, mentions that some expressions, phrases and sayings are always negative. Among them he finds professed expressions, like *EI OLE HULLUMPI/HASSUMPI* lit. ‘NEG

bepRES ACT NEG madSG CMP/sillySG CMP, id. ‘not too bad’, *EI OLE POIS TIELTÄ* lit. ‘NEG bepRES ACT NEG away roadSG ABL’, id. ‘it won’t be bad, I wouldn’t mind’, and expressions of indifference, like *EI TÄMÄN/TUON TAIVAALLISTA* lit. ‘NEG thisGEN/thatGEN divinesG PTV’, id. ‘not at all’. Sadeniemi (1946) notes that they have become permanently fixed into expressions with like signs. According to Sadeniemi (ibid 286-287), phrases like *EI MAISTU HONGALTA EIKÄ HAAVALTA* lit. ‘NEG tastepRES ACT NEG pine SG ABL NEGka aspensG ABL, id. ‘there is no taste at all’, *EI TÄSSÄ NAPILLA PELATA* lit. ‘NEG here button SG ADE playPRES PSS NEG’, id. ‘not being meticulous’, *X EI OLE EILISEN TEEREN POIKA* lit. ‘X NEG bepRES ACT NEG yesterdaySG GEN black grousesG GEN son’, id. ‘X is very experienced’, *EI OLLA JÄNIKSEN SELÄSSÄ* lit. ‘NEG bepRES PSS NEG haresG GEN backSG INE’, id. ‘there is no hurry’, etc. are all affective.

The object of Kiuru’s (1977) study is Finnish negopetal verbs. Negopetality (*kieltohakuisuus* in Finnish) is a phenomenon when a word, word form or phrase predominantly appears in a negative clause or other semantically negative contexts (Korhonen & Vilkuna 2005). Kiuru’s (1977) quantitative analysis based on dialectal data allows her to assign 9 different Finnish verbs a grade of negopetality which is a percentage of negative forms out of total number of examples for each verb (Kiuru 1977: 53). In the preface to her book Kiuru (1977) lists a number of synonymic idiomatic expressions meaning ‘nothing at all’, ‘not at all’, ‘never’, ‘nowhere’. According to Kiuru (ibid. 4) they belong to the lexicon of negative contexts, which is “affective in its definiteness”, e.g. *EI MAILLA EIKÄ HALMEILLA* lit. ‘NEG land PL ADE NEGka [burnt field]PL ADE’, id. ‘nowhere’.

Nenonen’s (2002) study of idiomaticity in Finnish emphasizes the relations between idioms and lexicon. In the chapter containing a description of Finnish verbal idioms she introduces a separate group of expressions that occur exclusively in the negative form. The proportion of such expressions among all verbal idioms is, according to Nenonen’s (2002) estimates, 46/1282, i.e. 4%. Nenonen (2002) says that their word order is typically frozen and it is generally impossible to change their form into positive without them losing their idiomatic meaning. She remarks (ibid. 62) that these expressions are not completely frozen, as it is possible to vary their tempus and insert adverbs into their structure. Among the class of other idioms Nenonen (2002) ranks such intensifying expressions as *X EI KISKO HOHTIMILLAKAAN TIETOA Y:STÄ ULOS*



lit. 'X NEG pull<sub>PRES ACT NEG</sub> pincers<sub>SPL ADE kAA</sub> information<sub>SG PTV YELA</sub> out', id. 'Y wouldn't tell X anything'; X *EI OTTAISI Y:TÄ EDES ILMAISEKSI/MAKSUSTA* lit. 'X NEG take<sub>COND ACT NEG</sub> Y<sub>PTV</sub> even free of charge/for money', id. 'X would never take Y'; exclamations like *EI MITÄÄN MUTTIA!* lit. 'NEG any<sub>PTV</sub> but<sub>PL PTV</sub>', id. 'No excuses!', *EI TULE MITÄÄN!* lit. 'NEG come<sub>PRES ACT NEG</sub> anything<sub>PTV</sub>', id. 'It will not succeed!'; parentheses, e.g. *EI MILLÄÄN PAHALLA MUTTA ...* lit. 'NEG any<sub>ADE</sub> evil<sub>SG ADE</sub> but ...', id. 'No offence intended but ...'; *EI TÄSSÄ NYT MUU AUTA KUIN ...* lit. 'NEG here now other help<sub>PRES ACT NEG</sub> than ...', id. 'We have no choice but ...' and other constructions (ibid. 64–65). Nenonen (2002) does not speak about negopetality in idioms as such, although on one occasion she mentions that *OLLA VÄLIÄ* lit 'be difference<sub>SG PART</sub>', id. 'to matter' is negopetal, i.e. it occurs generally in the negative form or otherwise in "contexts with a negative tinge" (ibid. 55).

Summarizing the above, one can see that among a relatively small number of studies that somehow touch upon the problem of negation in phraseological units (PUs) the leading tendency is to focus interest on negative polarity, i.e. the ability of certain PUs to appear exclusively (or predominantly) in the negative form (e.g. Martinez 1999). So far practically no effort has been made to give a valid description of negative modality in PUs and its influence on their negative polarity. In Chapter 5, Section 5.3.2 of the present book I attempt to fill this gap. Following Jespersen (1924) and Kiuru (1977) I do not consider negative and positive polarity merely as a matter of two extremes. Between them there is a gradation of positivity and negativity that different PUs can have. Resorting to quantitative analysis of large corpora can help one to place a PU on this scale, while an explicit formal representation of a PU's conceptual structure is needed if one wants to link the modality to the rest of the conceptual knowledge about this PU, on the one hand, and to those parts of discourse that linguistically encode it, on the other.

## 2.6 Summary

In this chapter, I have discussed several theoretical aspects related to the representation of semantic structure of phraseological units. An overview and criticism of previous approaches to these problems, both in the Russian theory of phraseology and western tradition, was presented. For a long time Russian scholars working in the field of phraseology have

primarily concentrated on comparing the structure of phraseological meaning to the structure of lexical meaning. At the same time, there has been no clear idea of what the structure of lexical meaning actually is, nor are there any valid formal methods for its description. The weakest point of these comparative methods, including recent multi-component models, was and remains the inconsistency between the scholars' firm conviction of the allegedly incontrovertible, but objectively rather vague, idea of lexical meaning and its structure as a sort of realistic semantic standard to which the phraseological meaning can be compared on the one hand, and a certain hypothetic character of their interpretation of the nature of meaning on the other (A. Žukov 1999), let alone their lack of formal devices for performing an explicit semantic analysis. As a result, the question of wherein lies the specificity of phraseological meaning still remains unanswered. The main problem of this approach to semantic structure and the reason why the latter cannot be equated with dictionary definition is obvious: as long as one strives for a psychologically plausible semantic theory, one needs such a model of semantic representation that is "considerably richer in both formal and substantive respects" (Jackendoff 1983: 3).

Following Nikanne's (1990) monostratal theory of representations, I assume only one level of conceptual representation. Consequently, the assumption of a special status of phraseological meaning as a different kind of meaning is simply impossible within the framework of Conceptual Semantics, which is used as the primary theoretical base of the present study. The relevant aspects of phrasal semantics cannot be determined simply by summing up the word meanings without the syntactic structure, and this is only in the simplest of cases. In more elaborate cases, it has to incorporate contextual elements as well. According to Jackendoff (2002: 333), the contextualized understanding of a sentence must be built from the following sources:

- The meanings of its words (lexical concepts)
- Conceptual structure conveyed by the grammatical structure of the sentence.
- Overall conditions on composed conceptual structure (well-formedness, plausibility).
- Conceptual structure derived from context.

In the present approach, the formal (i.e. formalized) structure of the expression has to incorporate all relevant aspects of information. Within

this formalized structure one has to look for correspondences between its different levels (tiers): the expression's syntactic configuration, conceptual structure (thematic arguments, functions, semantic field and modal tier), the referential tier and lexical components, which license their own correspondences between phonological, syntactic and conceptual information. If one assumes that these correspondences indeed exist, one has to show explicitly between what levels of the expression's formal structure they occur, by what means they are established, and what exactly corresponds or does not correspond to what. As Jackendoff (2002: 13-14) points out, correspondences mostly occur between composite units, not primitive elements of any of the levels. On the other hand, the units that are connected between phonology and syntax are not always the same units that are connected between syntax and conceptual structure.

I believe that the best way to solve this problem is to treat composite units, such as lexemes, PUs or constructions, as multi-level representations, and for this purpose I am going to use the Tiernet model proposed by Nikanne (2002, 2008a, 2008b). In the following Chapter 3, I will present an outline of the model and develop its application for the formal analysis of PUs. Since the model exists within the framework of conceptual-semantic theory, a few words have to be said about the theory itself.

### **3. Formalized representation of PUs in the conceptual-semantic Tiernet model – theoretical and methodological aspects**

In Sections 3.1–3.2 of this chapter I will briefly present the theoretical background of Conceptual Semantics and the Tiernet model. Section 3.3 reviews treatment of idioms and constructions within the framework of Conceptual Semantics (Jackendoff, Nikanne). The main objective of subsequent Section 3.4 is to demonstrate how different levels of representation and linking between them build up the internal structure of PUs.

#### **3.1 Conceptual Semantics and the Tiernet model**

Conceptual Semantics emphasizes the role of semantics as a connecting link between the theory of language and the theories of other cognitive domains (vision, social understanding, motor control etc.). Within the framework of Conceptual Semantics, the term concept is operationally used to mean “a mental representation that can serve as the meaning of a linguistic expression” (Jackendoff 1990: 11). The general aim of Conceptual Semantics is to create a psychologically plausible formal framework which can be applied to explanation of meaning and structure of concepts, “a formal syntax of concepts [...] that is in some ways much richer and in some ways much more constrained than standard logic” (Jackendoff 1990), and to develop an integrated and explicit model of language description where semantics is compatible with other cognitive faculties.

An important feature of Conceptual Semantics is the assumption of Representational Modularity of language and the human mind in general (Jackendoff 1987, 1992, 1997). According to this hypothesis, the human brain encodes information that is presented in different formats of representation. Each of these representational formats is a formal combinatorial system built out of its own set of primitives and principles of combination. Thus e.g. phonological, syntactic and conceptual structures are distinct representations which belong to separate modules (representational modules) and are linked together by principles of correspondence. Apart from representational modules there are also

mapping modules<sup>93</sup> (lexicon, morphology, constructions, DA system), which do not include any representations of their own, but instead license mapping between different representations of representational modules (Nikanne 2005a: 192-197). Nikanne's idea of modularity differs from that of Jackendoff.

In Jackendoff's model representations split up into different tiers. Jackendoff's (1997) hypothesis of representational modularity treats phonology, syntax and conceptual structure as autonomous modules, which have their own primitives and combinatorial principles and which license their own well-formed levels of representation. A representation is autonomous if it cannot be reduced to another level of representation (Nikanne 2008a, 2008b). But representations in their turn can be seen as a combination of different tiers, which are independent parts of their representations. In his latest works, Jackendoff (2002) motivates dividing conceptual structure into tiers, each tier conveying a different aspect of sentence meaning.

As Nikanne (2005) points out, it is impossible to know whether an autonomous structure is a representation of its own, or if it is a tier of some other representation. In Nikanne's theory, the necessity of Jackendovian representations is questioned. Instead the organization of grammar (which Nikanne calls the *Tiernet* model) is based on autonomous tiers and linking between them. The idea for such organization of language was derived from autosegmental phonology (Goldsmith 1979; 1990; Liberman & Prince 1977), which divided phonological structure into a number of tiers. Any formally independent sub-system which has its own primitives and principles of combination can form its own tier or micro-module. Since micro-modular representations are irreducible to over levels, they can be regarded as truly autonomous. The main idea is to keep micro-modules very simple and instead put the main emphasis on the linking principles that govern the correspondences between them. Figure 26 below shows the *Tiernet* organization of the Finnish grammar suggested by Nikanne (2002, 2006, 2008a, 2008b).

---

<sup>93</sup> Jackendoff (1990: 155) calls them correspondence rules.

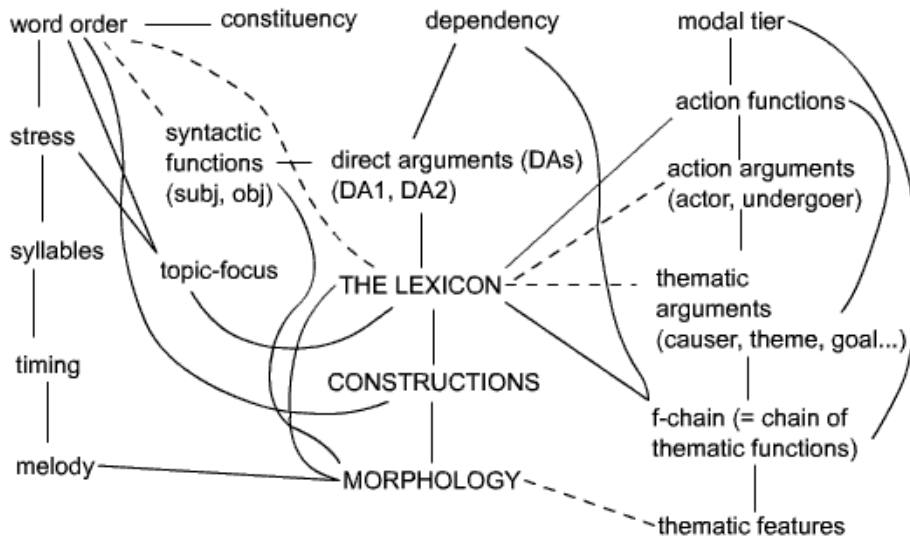


Figure 26 Finnish grammar represented in a Tiernet model (Nikanne 2002, 2006, 2008a, 2008b)

Nikanne argues that a micro-modular model has several important advantages (Nikanne 2002):

- It is open and dynamic. Languagewise differences may be explained by the absence of certain tiers and differences in linking across tiers in different languages.
- There is no need to ask what belongs to syntax and what to semantics, since the model breaks traditional borders of language faculties.
- We are dealing with very simple structures.
- Besides being modular, the model also has connectionist features, e.g. it assumes that the strength of linking can be weaker or stronger between different tiers.

In the current chapter and in Chapter 5 of this book I will adapt and develop descriptive tools of Conceptual Semantics and in combination with the Tiernet-inspired way of representing language structures apply those to the task of formal PU analysis. It is my major goal here to develop a new model of PU representation, which will be compatible with theoretical assumptions and methodological guidelines of Conceptual Semantics. For a more complete overview of general assumptions and methodological guidelines of Conceptual Semantics see Nikanne (2008a).

## 3.2 Organization of the conceptual structure and lexical linking rules

According to The Conceptual Structure Hypothesis, conceptual structure (CS) is a single level of mental representation that encodes linguistic meaning and onto which and from which all peripheral information is mapped (Jackendoff 1983: 17, 19; 1995: 138). It is the form in which speakers encode their construal of the world and at which linguistic, sensory and motor information are compatible. Lexical Conceptual Structure (LCS) is an abstraction with language-independent properties that goes beyond structural idiosyncrasies (Jackendoff 1983, Jackendoff 1990, Jackendoff 1996). According to Jackendoff (1990: 155), a theory of conceptual structure has to be able to express the semantic distinctions among sentences and the inferences that sentences support.

### 3.2.1 Linking types

Jackendoff notates correspondences between phonological, syntactic and conceptual structures with a system of subscripts: phonology-syntax correspondences are marked with pre-subscripts and syntax-semantics correspondences with post-subscripts. Nikanne (2005a: 192) uses the terms *correspondence*, *mapping* and *linking* as synonyms. In the present approach, the difference will be made between *linking* as a general term for all types of connections that exist in the Tiernet between its tiers, and the *correspondence link* as a special type of linking. This differentiation relies on the fact that not all relations between tiers are based on correspondence. For instance, the well-formedness principles of conceptual structure are to a large extent based on selection (dependency) relations that exist between its zones and tiers (Nikanne 2005a: 201): *f-chain* is a dependency chain where relations can select each other, the *f-chain* functions and the *act-chain* functions select thematic arguments, the *semantic field* is a tier which selects particular functions as its immediate scope, the *modal tier* functions can select the *f-scheme* functions as well as each other as their immediate scope. Some of these links can be seen e.g. in Figure 27 below, which shows the organization of conceptual structure as proposed by Nikanne (2005). The three main tiers here are the *f-chain*, the *argument level* and the *act*

chain. They will be described in more detail in the following Sections 3.2.2–3.2.4.

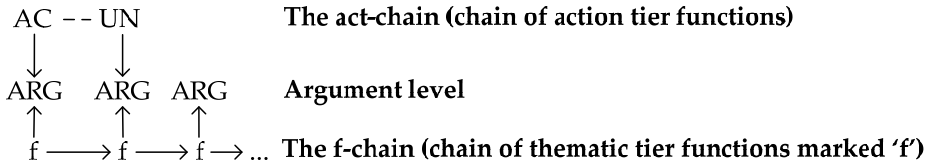


Figure 27 Organization of conceptual structure (Nikanne 2005a: 202)

Phrasal syntactic tree structures are connected by constituency links. As Jackendoff observes (2002: 52), the lexical formation rules (morphology) include tree fragments from which such structures can be built, i.e. they are also based on constituency. Syntactico-semantic and phonology-syntax linking is governed by correspondence rules (Jackendoff 1997: 24). In addition to the above-mentioned, there may be some other types of linking, such as assignment (e.g. the function TO assigns the thematic role GOAL to its argument). In the present book, the distinction will be made between at least three types of linking relations: constituency, selection (dependency) and correspondence. The following notation is used for these links (Figure 28 below):

- Correspondence is marked with a dotted line.
- Selection is marked with a solid arrow pointing at the selected item.
- Constituency is represented by a solid line with a round dot pointing at the constituent node.

**Correspondence**       $X \cdots Y$       (X corresponds to Y)

**Selection, dependency**  $X \rightarrow Y$       (X selects Y, X is dependent on Y)

**Constituency**       $X \rightarrow \bullet Y$       (Y is a constituent of X)

Figure 28 Linking types

### 3.2.2 Thematic tier

Conceptual structure operates with conceptual categories, such as SITUATION, PLACE, PATH, THING, PROPERTY, AMOUNT etc. Nikanne (2000: 80) divides them into two types: simple and complex. Simple categories (THING, PROPERTY, AMOUNT etc.) are not governed by event structure functions and are typically in argument positions.



Complex categories (SITUATION, PLACE and PATH) are complexes of functions and their arguments; they must be governed by event structure functions (Nikanne 2000: 80-81). Jackendoff (1987, 1990) and Nikanne (1990b, 2000, and 2005a) assume several tiers within the conceptual structure representation of the category SITUATION: the thematic tier, the action tier, the temporal tier.

The thematic tier encodes aspects of the situation structure that have to do with change or state of affairs or causation (Nikanne 2008a). Nikanne (1990b) divides the thematic tier functions into three groups called zones according to their scope: zone 3 includes causative and inchoative relations, zone 2 non-causative Event or State relations and zone 1 place and path relations. Figure 29 below is taken from Nikanne (1997a, 2005a). It shows the thematic tier functions, the thematic roles and the structure of each zone.

<b>Zone 3</b> <b>Causative and inchoative relations</b>	<b>Zone 2</b> <b>Non-causative Event or State relations</b>	<b>Zone 1</b> <b>Place and path relations</b>
<i>Non-monadic functions:</i> CAUSE	<i>Non-monadic functions:</i> BE, GO, STAY, ORIENT, EXT	<i>Non-monadic functions:</i> (none)
<i>Monadic functions:</i> INCH	<i>Monadic functions:</i> CONF, MOVE	<i>Monadic functions:</i> AT, IN, ON, UNDER, TO, TOWARD, FROM, AWAY-FROM, VIA
<i>Thematic role:</i> Causer, Agent	<i>Thematic role:</i> Theme	<i>Thematic role:</i> Landmark (Location, Goal, Route, Source)
<i>Structure:</i> <Causer>	<i>Structure:</i> Theme	<i>Structure:</i>
↑ f3 → SITUATION	↑ f2 → <[PLACE/PATH]>	f1 → {PLACE/PATH/Landmark}

Figure 29 The thematic tier zones and their functions (Nikanne 1997a, 2005a)

Monadic functions can select only one complement, which can be either a thematic argument or another function. Non-monadic functions can select more than one complement. Nikanne (1990b, 2006) marks the monadic root nodes as **F'** and non-monadic as **F**. The lower case **f** marks nodes where the distinction between monadicity vs. non-monadicity is not essential. The backbone of the conceptual structure is a chain of functions, or the f-chain, which follows the well-formedness principle called the f-schema:  $f3^* \rightarrow f2 \rightarrow f1^*$ . Numbers 1, 2 and 3 indicate the three zones. Well-formedness is based on dependency (selection)

indicated by an arrow. An asterisk indicates that there can be zero or more functions of this zone in the chain (Nikanne 1997a, 2002, 2005a).

According to Nikanne (1990b, 2006), Jackendoff's thematic functions can be analyzed further as feature hierarchies. These features are presented in Table 3 below.

*Table 3 Thematic features (Nikanne 2006: 221-222)*

Abbreviation	Feature	Description
b	Bounded	Indicates boundedness of a path or event.
c	Contacted	Indicates that the theme/figure is in contact with the Landmark. Sub-feature of b.
a	Attached	Indicates that the theme/figure is attached to the Landmark. Sub-feature of c.
D	Directed	Indicates that the event has a direction.
3D	3 dimensional	Sub-feature of b.
in on under	in on under	Place features indicating a place in, on or under the Landmark. Sub-features of 3D.
gl	Goal	Path feature indicating the end of the path. Sub-feature of D.
so	Source	Path feature indicating the beginning of the path. Sub-feature of D.
ro	Route	Path feature indicating a relevant midpoint on the Path. Sub-feature of D.
di	Distributed	Indicates that the theme/figure is distributed over the Path or Place.
T	Time-related	Indicates that the event has an internal time line.

Thus, each function can be represented as a root node plus a combination of relevant features, e.g. TO as  $[F'1 [b][D [gl]]]$ , GO as  $[F2 [T][D]]$  and CAUSE as  $[F3 [T][D] \rightarrow F'3 [b][D [gl]]]$  (Nikanne 1990b, 2006).

### 3.2.3 Thematic arguments (theta-arguments)

Thematic arguments are not considered as primitives. They are elements selected by thematic tier functions (f-chain) according to the following Argument Selection principles (Nikanne 2005a: 202):

- Each F must select a thematic argument.

- Each f2 must select a thematic argument.
- No f can select more than one thematic argument.

Thematic arguments are assigned different roles depending on the thematic function which they are selected by. Thus, the roles of Causer and Agent are assigned to the theta-argument selected by an f3, the role of Theme is assigned to the theta-argument selected by an f2 and the role of Landmark to the theta-argument selected by an f1. The latter role can vary between Location, Goal, Route or Source, depending on the sub-features of D, carried by the root node.

### 3.2.4 Action tier

Adapting Culicover & Wilkins (1986) and Talmy (1985), Jackendoff (1990) proposes an action tier, which encodes dominance relations between participants of a situation and basically corresponds to roles in human social relations. Jackendoff (1990) assumes two relational roles assigned by the action tier – Actor and Patient (or Undergoer), i.e. an active and a passive participant of a situation. He describes them as arguments of the function AFF (affect). The first argument of AFF is the Actor; the second is the Patient (Undergoer) (Jackendoff 1990: 127). In order to distinguish between a wilful and a non-wilful doer, Jackendoff (1990) introduces a feature elaboration [ $\pm$ volitional] or [ $\pm$ vol]: a volitional Actor is the first argument of AFF<sub>+vol</sub>, and a non-volitional Actor is the first argument of AFF<sub>-vol</sub>.

The effect experienced by the Undergoer can be positive (UN+ selects a benefactive Undergoer, or Beneficiary), negative (UN– selects a malefactive Undergoer, or Patient) or neutral (UN selects a neutral, unmarked Undergoer) (Nikanne 1995). Nikanne (1995: 7-8) uses the following tests to assign AC and UN roles for Finnish:

- |         |   |
|---------|---|
| Test 1. | Se mitä X tekee on S. 'What X does is S' => X = Actor                           |
| Test 2. | Se mitä X tekee Y:lle on S. 'What X does to Y is S' => X = Actor, Y = Undergoer |
| Test 3. | Se mitä Y:lle käy/tapahtuu on S. 'What happened to Y is S' => Y = Undergoer     |

In Nikanne's (1995) version of the action tier both AC (Actor) and UN (Undergoer) roles are functions of the form G(Arg), where 'G' is an unspecified action tier function and 'Arg' is its argument. He refers to them with the term act-functions, where act stands for 'action tier'. Together they form a unit called the act-chain, i.e. the chain of action

tier functions, which select Actor and Undergoer. Nikanne's (2005) notation can be found in Figure 27 (Section 3.2.1). A dashed connector line between AC and UN is meant to express their belonging to the same complex unit. However, the same relationship could also be conveyed by assuming a function tier ACT (similar to Jackendoff's AFF function), which selects its arguments. The dominance relation between the roles of the action tier can e.g. be described by correspondences to rank features of a (social) hierarchy Dominance dimension – AC corresponds to a dominating Superior rank (or [sup]) and UN to a dominated Subordinate rank (or [sub]) (Figure 30 below).

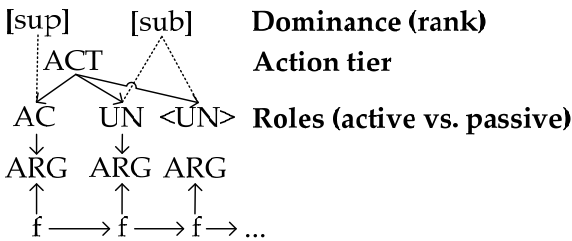


Figure 30 Action tier

Benefactory vs. malefactory Undergoer can be formalized with the help of an Effect on well-being level with features [benef] and [malef]. For example, the [benef] feature will select an Undergoer in the verb *auttaa* 'help', and the [malef] feature an Undergoer of the verb *kärsiä* 'suffer' (Figure 31 below).

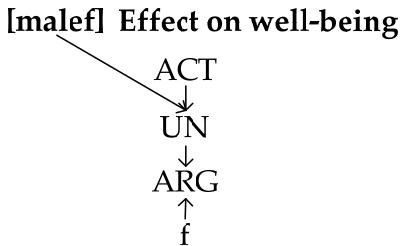


Figure 31 Action tier of the verb KÄRSIÄ 'suffer'

In a similar way a wilful vs. non-wilful doer can be formalized by adding a Volitionality level with features [+vol] and [-vol], which select a suitable Actor or Undergoer. Figure 32 below describes the action tier of the Finnish verb *tyrkyttää* 'force on', where Recipient is a non-wilful Undergoer.

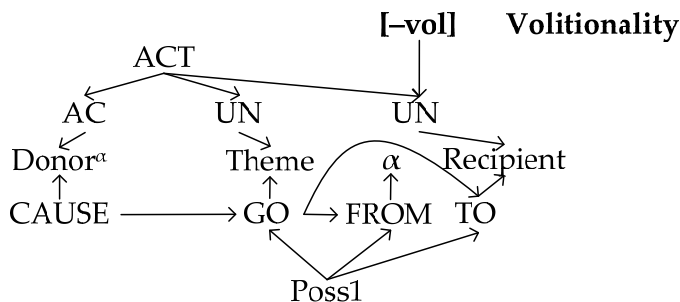


Figure 32 Action tier of the verb *TYRKYTTÄÄ* 'force'

The general principles of action tier argument selection are formulated by Nikanne (1995: 10) as follows:

- (A) No actors in zone 1. I.e. no actors are Locations, Goals, Sources or Routes.
- (B) The same argument cannot be selected by more than one AC.
- (C) An Actor must cs-command an Undergoer selected by the same act-chain.

Cs-command is a term introduced in Jackendoff (1992b). Cs-command is defined in Nikanne (1995: 10) as follows:

X and Y are arguments of the same f-chain. X cs-commands Y if X is a complement of a function that has scope over the function whose complement Y is.

The left to right order in f-chain establishes a thematic hierarchy and AC falls higher in this hierarchy than UN, which results in the principle according to which AC always selects the leftmost argument of the lexical f-chain.

### 3.2.5 Temporal tier

According to Jackendoff (1990), the internal time flow of the situation is encoded in the separate tier, which is called the temporal tier (T-tier). It has two primitives: point of time ('P' or 'I') and region of time ('R' or '-----'). R stands for the timeline itself, while P is a possible boundary of R. Region of time is directed from left to right, although it is omitted from most notations. Further development of the theory of temporal tier can be found in Nikanne (1990b, 1997b and 2002) and Pörn (2005). According to the temporal tier licensing principles presented in Nikanne (1990b: 179) R is required if there is a T feature in the thematic tier. P is required at the beginning of R if there are [so] and [b] features, at the end of R if there are [gl] and [b] features and inside R if there are [ro] and [b] features. P and R can appear alone as P, R, or they can be

combined as PR (|----), RP (----|), PRP (|----|) and RPR (----|----) (Nikanne 1990b, 2008a). Pörn (2005) uses subscripts after P in order to distinguish between the starting point ( $P_s$ ) and final point ( $P_f$ ) of a situation. Following the same logic, one can introduce a  $P_i$  to indicate the intermediate point of time.

In Nikanne (1997b) another part of the T-tier is introduced, the CT-tier ('C' as in 'construction', 'common' or 'complex'). It is a schematic temporal tier that corresponds to the linear time flow and is shared by all situations expressed in a complex sentence. The individual T-tiers of the situations expressed in the CS are separately related to the CT-tier in three following ways: they can either be equal to ( $=$ ), included in ( $\in$ ) or not included in ( $\notin$ ) the CT-tier. The CT-tier can either be undivided, or it can be divided into two parts (CT1 and CT2) so that its first part (CT1) represents the earlier period of time and the subsequent part (CT2) the later period of time. Nikanne (1997b) leaves open the question of whether the CT-tier can include more than two parts. He admits that this is a possibility, the other possibility being that the CT-tier always connects only two situations at a time and several such CT-tiers can be then connected to each other. Nikanne (1997b) represents the undivided and divided CT-tiers as follows:

- Undivided CT-tier:

CT

-----

- CT-tier divided into two parts:

CT1    CT2

-----|-----

In Figure 33 below I attempt to combine the thematic tier and the T-tier into the same network structure. This picture is an elaboration on the description of the situation *W went from X to Y via Z* presented in Nikanne (1990b: 178). The temporal structure of this situation includes several levels:

- The level of primitives (R, P)
- The level of temporal features: [s] stands for 'starting', [i] for 'intermediate' and [f] for 'final'
- The T-tier proper
- The CT-tier

Primitives and their features are connected to the thematic structure in accordance with the temporal tier licensing principles – a slight modification of those presented in Nikanne (1990b: 179):

- T licenses R
- [b] licenses P
- [so] licenses [s]
- [ro] licenses [i]
- [go] licenses [f]

Thus, each temporal feature is connected by a correspondence link to a thematic feature which licenses it: R corresponds to T, P corresponds to [b], [s] corresponds to [so], [i] to [ro] and [f] to [go]. Temporal features [s], [i] and [f] are connected to the respective point of time primitive by means of selection links, i.e. a starting point of time is a P which is in the scope of the feature [s].

The T-tier proper is connected to its primitives by constituency links. In Figure 33 it has a form of  $P_s R P_i R P_f$  and is constituted by the region of time (R), which is the time line itself, and three boundaries: the starting point of time ( $P_s$ ), the intermediate point of time ( $P_i$ ) and the final point of time ( $P_f$ ). Finally, the T-tier is connected by a correspondence link to the CT-tier. This link represents the same relation which Nikanne (1997b) describes as  $X = Y$  ('X is equal to Y'). In other words, in this particular temporal structure, the T-tier is equal to the CT-tier.

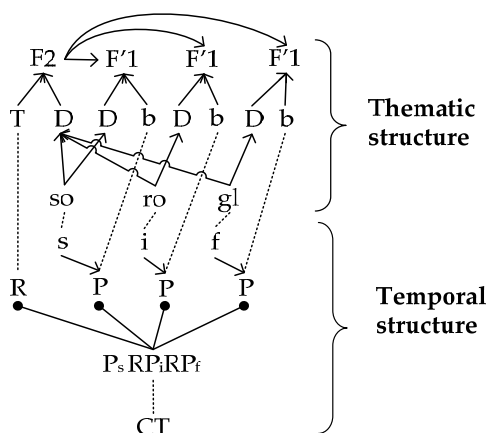


Figure 33 Temporal structure and its connection to the thematic structure represented as a network

Figure 34 and Figure 35 below present an attempt to apply the network model to the analysis of two types of associations of the causative

functions. Jackendoff (1990) distinguished two possible cases – launching (Fin. *sysäys*) and entrainment (Fin. *saatto*). Pörn (2005) adds a third type – *aavistus* ‘foreboding’.

Launching represented in Figure 34 is the kind of causation seen e.g. with the Finnish verb *heittää* ‘throw’. The causation is associated only in the starting point of the temporal tier of the caused Event. The end boundary of the temporal tier of zone 3 is the starting boundary of the temporal tier of the core zones (Nikanne 1990b: 188-190). Nikanne (1990b: 189) indicates this correspondence with a colon placed between the two equated boundaries.

Z3:            ----|  
                 :  
Z2&Z1:        |----

The network model required some adjustments to this notation. Thus, both temporal phrases ( $RP_f$  and  $P_sRP_f$ ) in Figure 34 include the same primitive constituent  $P$ , which is at the same time the final point of time in the temporal tier of zone 3 ( $RP_f$ ) and the starting point of time in the temporal tier of zones 2 and 1 ( $P_sRP_f$ ). This is indicated by letting  $P$  be selected by both  $[f]$  and  $[s]$  features. The CT-tier is divided into two parts – CT1 and CT2. The T-tier of zone 3 corresponds to CT1 and the T-tier of zones 2 and 1 corresponds to CT2. From the chronological point of view, this means that causation happens earlier than the caused event.

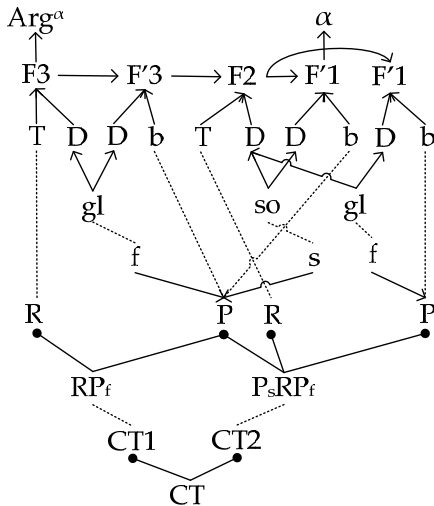


Figure 34 Launching



Entrainment (Figure 35) is a relation where causation and the caused Event are temporally coextensive, i.e. the former lasts as long as the latter. This kind of causation is seen e.g. in the Finnish verbs *vetää* 'drag' or *tuoda* 'bring'. Nikanne (1990b: 188) suggests that one can understand entrainment so that the T-tier of zone 3 is equal to the T-tier of the core zones. While Nikanne uses the notation "=", I attempt to represent the same relationship by means of correspondence links. In Figure 35 both temporal phrases correspond to the same CT-tier, which means that they are simultaneous.

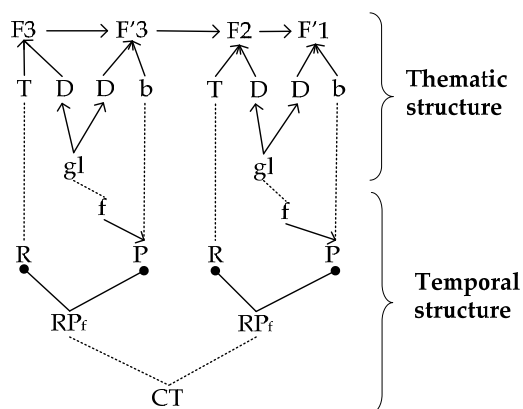


Figure 35 Entrainment

### 3.2.6 The semantic field tier (S-tier)

Conceptual Semantics reveals conceptual parallelism in the same patterns which can be used to describe physical objects in space and for non-spatial cognitive domains such as possession, spatial location, time etc. While cognitive linguists tend to handle such cross-domain parallelisms as derivational (cf. Lakoff & Johnson's 1980 view on metaphor), Jackendoff argues that they are instantiations of a more abstract schema – the same semantic function (Jackendoff 1976, 1983, 1992, 2002: 359) which can be specialized by means of a feature called semantic field. Conceptual Semantics promotes the localist idea of Gruber (1965), which ascribes the primary value to the semantics of motion and location (i.e. spatial semantics) and states that over fields can be derived from it by means of rewriting rules (see e.g. Jackendoff 1983). In Nikanne's (1990b) hierarchy of semantic tier fields, the spatial semantic field has the highest rank since it can be applied to all thematic functions.

Nikanne (2002) distributes semantic fields between three zones of the thematic tier. Semantic fields of zone 1 (Place and Path relations) and zone 2 (non-causative Event or State relations) are e.g. spatial (Spat), possessive (Poss), temporal (Temp), circumstantial (Circ), characterizing (Char) and existential (Exist). These two zones must belong to the scope of the same field. Semantic fields of zone 3 (causative and inchoative relations) are physical (Phys), social (Soc), magic (Mag) etc. Nikanne admits that semantic fields of zone 3 are to a large extent independent of those that select zones 1 and 2 as their scope. For more discussion on semantic fields see Nikanne (1990b, 2002). Inasmuch as semantic fields specify the cognitive domain to which the described situation has to be placed (Nikanne 2002) the features of this particular domain are applicable to the situation. Some lexical items can encode certain distinctive features, while others, belonging to the same domain, leave them unspecified, cf. the description of Dutch verbs that refer to curvative distinctions given in van der Zee (2000) or the grain levels in the linguistic expressions of motion in van der Zee, Nikanne & Sassenberg (2010). The most explored domain in Conceptual Semantics is the spatial field.

In Leino et al. (1990) different uses of Finnish local cases are examined through the prism of semantic fields. Nikanne's (1990a) contribution concerns the possessive field. As Nikanne points out, the field of possessivity is not limited to alienable (Nikanne marks it with the symbol Poss1) and inalienable possession (Poss2), but also includes cases where an inalienable possession cannot be regarded as an integral part of the possessor, e.g. the flu in the sentence *Pojalla on flunssa* 'The boy has a flu' (Poss 3). Nikanne (1990a) also singles out cases where a SITUATION or an EVENT belongs to something/somebody, e.g. *Michael Jacksonin kuolema oli suuri menetys maailmalle* 'Michael Jackson's death was a big loss for the world' (Poss 4). Finally there is group marked as PossM, which includes cases of mental and communicative possession, e.g. *Minulla on ajatus* 'I have an idea', *Sain uutta tietoa Pekalta* 'I got new information from Pekka' etc.

Figure 36 and Figure 37 below demonstrate how two Finnish verbs with identical f-chains differ in the semantic field: the verb *antaa* 'give' specifies a Possessive (Poss1) field in zones 2 and 1, while in the verb *lähettää* 'send' thematic functions of these zones are selected by the Spatial field and the zone 3 function is in the scope of the Physical semantic field.

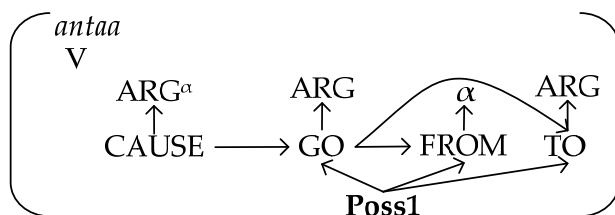


Figure 36 Possessive semantic field in the Finnish verb *ANTAA* 'give'

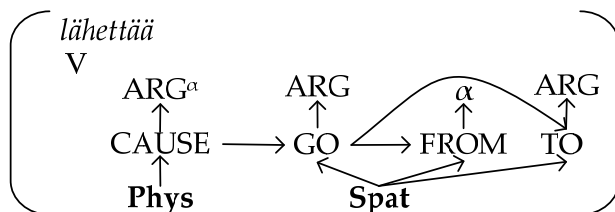


Figure 37 Physical and spatial semantic fields in the Finnish verb *LÄHETTÄÄ* 'send'

I will return to the topic of semantic fields in Section 3.4.3.6.1 of the current chapter with regard to semantic analysis of PUs.

### 3.2.7 The modal tier

As Nikanne (2002) points out, the modal tier is not described in any literature on Conceptual Semantics, but modality itself is a well studied area in linguistics. Nikanne (2002) argues that LCS has to include the modal tier, which is needed to describe the probability, possibility, negation etc. of the situation or its parts expressed by the proposition. The modal tier is a separate part of the LCS and its primitives are e.g. features like *Prb1* (probability), *Pot* (potentiality), *Cert* (certainty), *Neg* (negation), etc. This featured can be expressed in language by various means: modal verbs and adverbs, infinitive constructions etc. The modal tier features are linked to the thematic tier functions and to each other by means of selection ( $\rightarrow$ ), i.e. a feature selects some thematic function or other modal feature as its immediate scope. For instance:

(63) *Myös liian runsas kastelu voi aiheuttaa lehtien kellastumista.*

'Too abundant watering as well **may** cause yellowing of leaves.'<sup>94</sup>

<sup>94</sup> <http://www.huonekasvit.net/foorumi/viewtopic.php?p=384#p384>, Feb 3 2006

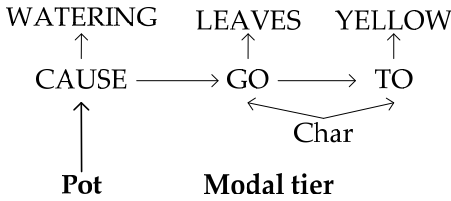


Figure 38 Modal tier in Example (63)

In (63) (Figure 38) the Pot feature chooses the CAUSE function as its scope, thus expressing that the causative situation as a whole is possible.

- (64) [Ilmastonmuutos **tuskin** aiheuttaa itsessään uusia, aiemmin tuntemattomia sairauksia]<sub>1</sub>, [mutta se **varmasti** lisää ihmiselle haitallisia sairausilmiöitä]<sub>2</sub>.  
 '[Climate change in itself will **hardly** cause any new, previously unknown diseases]<sub>1</sub>, [but it will **certainly** add disease phenomena harmful to people]<sub>2</sub>'.

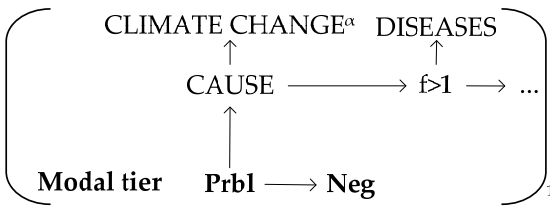


Figure 39 Modal tier in Example (64)

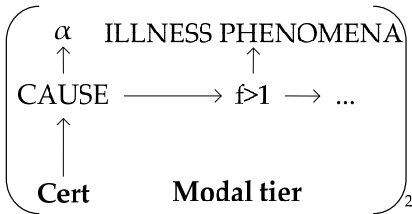


Figure 40 Modal tier in Example (64)

Figure 39 and Figure 40 above present an analysis of the modal tier in Example (64). The first conceptual clause in Figure 39 encodes the modality of non-probability of the causative situation: the Prbl feature selects both the CAUSE function and the Neg modal feature as its scope. The second conceptual clause in Figure 40 includes the modality of certainty: the thematic function CAUSE is selected by the modal feature Cert.

The modal tier in PUs will be examined in some more detail in Section 3.4.3.6.2 of the current chapter and in Section 5.3.2 of Chapter 5.

### 3.2.8 Principles of semantico-syntactic correspondence

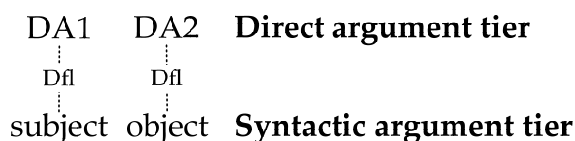
Nikanne (2005a: 203-205) suggests that there are several correspondence principles that exist between different levels of LCS. The default correspondence between the lexical f-chain level and the syntactic category level is that  $f > 1$  corresponds to V and  $f1$  corresponds to P. According to Nikanne, lexical conceptual structure determines which conceptual argument is linked to which syntactic argument. He suggests that theta-arguments (conceptual arguments) are not to be linked directly to syntactic arguments (subject and object). Instead there is a mediating subsystem between them that determines the syntactic roles of each theta-argument – the direct argument system (DA system). Direct argument (DA) stands for a word's syntactic argument which is not licensed by any adjunct rule or other structure specific linking rule (Nikanne 1997a: 87). DAs are determined by a lexical item, not in syntax. Default linking between conceptual structure and syntax (which basically corresponds to the rules of grammar) is constrained by the interaction between the f-chain and the DA system. Nikanne suggests the following two principles which can be used to determine which conceptual arguments may appear as syntactic arguments and which may not:

- If a function in the lexical f-chain requires a theta-argument, this theta-argument is a potential DA.
- If a theta-argument is marked as implicit ([...]<sup>I</sup>) in the LCS, it is not a potential DA (Nikanne 1997a: 88; 2005a: 204).

Nikanne assumes that at least in Finnish and English LCSs cannot have more than two DAs. Potential arguments are marked as DA1 and DA2 from left to right and thus form a hierarchy  $DA1 > DA2$ . They are linked to syntactic arguments according to the following default linking principle:

- By default DA1 corresponds to subject.
- By default DA2 corresponds to object.

Figure 41 below presents the default linking between the DA-tier and the syntactic argument tier (Nikanne 1997a, 2005a).



*Figure 41 Default linking between the DA-tier and the syntactic argument tier (Nikanne 1997a, 2005a)*

### 3.2.9 Context and reference – the referential tier

Jackendoff (2002b: 280-285) discusses several approaches to meaning, which address the question of whether there is a specifically linguistic part of semantics that can be distinguished from contextual meaning. The two ways such separation could be achieved is either by locating contextual meaning to a special level of structure connected to linguistic semantics by the pragmatic interface, or by assuming that linguistic semantics is a subset of contextual meaning. Jackendoff (2002b: 201, 208, 214) himself advocates a view that there is no special level of linguistic semantics except for conceptual structure. Thus, contextual information is integrated<sup>95</sup> with conceptual material derived from linguistic expression by the same conceptual integrative processor which binds the conceptual structure of a lexical item to the part of the thought being expressed. However, in order to establish the contextualized meaning of an utterance, one has to establish its reference (ibid. 324).

In a conceptualist theory, reference is dependent on the language user: a referent's conceptualization by the speaker is a necessary (but not sufficient) condition for a speaker to refer. Jackendoff formulates a conceptualist theory of reference as follows:

A speaker S of language L judges phrase P, uttered in context C, to refer to entity E in [the world as conceptualized by S] (Jackendoff 2002b: 304).

Basically this means that the speaker is referring to his mental representation of an entity, rather than to the entity itself. As Jackendoff points out, the standard approach to reference – assuming that linguistic expressions refer to objects out in the world – suffers from two substantial problems. On the one hand, if language is in the minds of language users, the connection between the human mind and the world outside the mind becomes unclear. On the other hand, the notion of objects in the world is not so self-evident either (Jackendoff 2002b: 295-303). As is shown in Figure 42 below, Jackendoff's conceptualist view denies any direct connection between the form of concepts and the outside world.

---

<sup>95</sup> Jackendoff (2002b: 210) describes integration as a part of lexical access. It is a process of combining an item in a working memory with a larger structure being built. "For each set of formation rules that defines a level of linguistic structure, the language processor requires an integrative process that uses these principles to construct structures at that level" (ibid. 198).

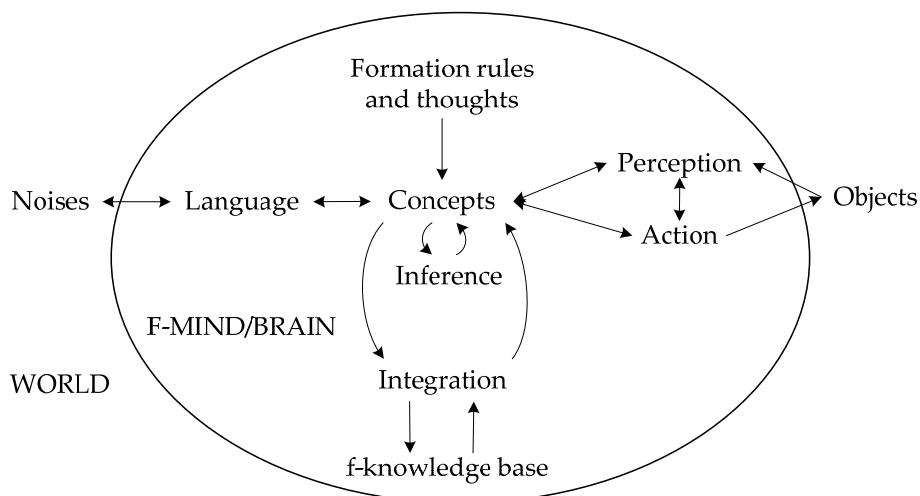


Figure 42 The conceptualist view (Jackendoff 2002b: 305)

The only contact the brain (and language) has with the world is via perception (through the complex mediation of the visual system) and action. In response to stimulation from the outside world, perceptual systems construct percepts – cognitive/neural structures that distinguish individuals in the perceived environment and permit one to attend to one or another of them. The experience that accompanies having a percept is that of an object in the world, although some of these objects may actually lack physical reality (Jackendoff 2002b: 307-309). A percept contains the following features:

- Descriptive features, e.g. the percept's shape, size, colour, part structure, location, motion, character of motion etc.
- An indexical feature (a percept's index) – a figural characteristic, which distinguishes a figure from the background and to which descriptive features can be attached.

A percept's index is what enables it to be tracked over time as it changes position and even properties; it is the index that makes a percept count as 'the same thing' with a history over time (Jackendoff 2007a: 102-103).

An index is what gives a percept its 'that-ness' – it is not just a collection of perceptual features, but an individual. Computationally, an index is what enables percepts in different modalities to be bound together... (Jackendoff 2007a: 103).

- One or more modalities in which descriptive features are present, e.g. visual, auditory and tactile modalities.

- A valuation which registers the status of the cognitive structure in a number of modality-independent dimensions: e.g. external (i.e. being experienced as ‘out in the world’) vs. internal (being experienced as an image), familiar vs. novel, self-produced vs. non-self produced, meaningful vs. non-meaningful, mattering vs. non-mattering (registers emotional effects produced by the object) (Jackendoff 2002b: 310-313).

According to Jackendoff (2002b: 311-315), the crucial feature for linguistic reference is the indexical feature of a percept: if there is no indexical feature, there is nothing to which a referring linguistic expression can be linked. Demonstratives (*this*, *that*) have minimal descriptive content and thus express mainly an indexical feature. A distinction between kinds versus instances can be made by assuming that the former lack an indexical feature in their CS, while the latter have one. An instance can be formed from a kind by adding an indexical feature and, conversely, a kind can be formed by deleting an indexical from an instance (Jackendoff 2002b: 319). Unlike perceivable entities, abstract objects and categories have conceptual structure which contains only inferential descriptive features and no connection to perceptual interfaces whatsoever. Still, one is able to refer to them as long as they have an indexical feature (Jackendoff 2002b: 323). Any referred entity can be also classified into different ontological types such as object (thing), action (event), location, sound, tactile sensation, manner, distance etc (Jackendoff 2002b: 316-317).

Yet another difference has to be made between the reference purported by the speaker, and the satisfied reference established by the hearer. According to Jackendoff, a referential expression succeeds in referring for the hearer if it is satisfied by something that can serve as its referent (Jackendoff 2002b: 324). The purported reference can be established in the hearer’s conceptualization of the world, if the purported referent is present in the hearer’s f-knowledge<sup>96</sup> base or the readily available context. If the purported referent is in conflict with the hearer’s f-knowledge base or the readily available context, or it contains descriptive features that inherently conflict with each other, the hearer has to fall back on some repair strategies (ibid. 325). The notion of satisfaction applies both to the reference of NPs and sentences, but sentences have an additional layer in

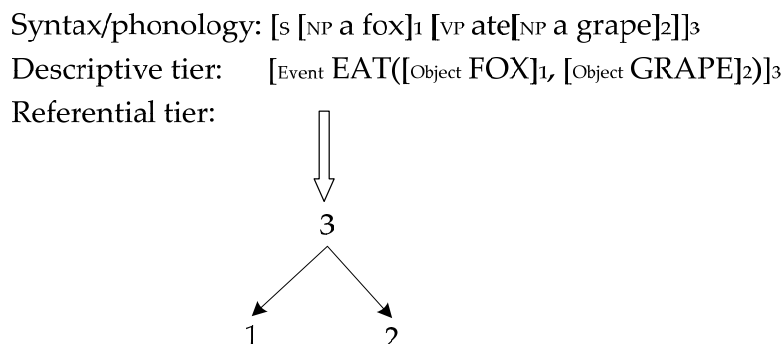
---

<sup>96</sup> Jackendoff (2007b) explains that the term f-knowledge or functional knowledge is introduced in Jackendoff (2002b) to distinguish the speaker’s-hearer’s knowledge of language from conscious knowledge of facts and from philosophers’ notion of justified true belief.



which they are characterized as true or false on the basis of how they are referentially satisfied (ibid. 327).

Furthermore, Jackendoff introduces a separate tier, which records active referents in the discourse, thus organizing the referential claims about the entities of the sentence. He calls it the referential tier (Jackendoff 2002b: 394ff). Figure 43 below presents Jackendoff's outline of the referential tier.



*Figure 43 Outline of the referential tier offered by Jackendoff (2002)*

In Figure 43, indices of the referential tier correlate with the two Object constituents and the Event constituent of the descriptive tier, and with constituents of the syntax/phonology. In terms of Jackendoff's theory of reference briefly described above, these indices are the indexical features invoked by the sentence: they pick out what constituents of the sentence are intended to correspond to individuals in the world as conceptualized by the speaker. The presence of the indices encodes the existential claims that go with the sentence: e.g. index 3 in Figure 43 is the claim that the event of the fox eating the grape took place (Jackendoff 2002b: 395). The arrows pointing from 3 to 1 and 2 represent referential dependence of the claimed existence of the event on the claimed existence of its participants. The block arrow pointing at the event's index represents the assertive force which allows us to infer that the event took place (Jackendoff 2002b: 399-400).

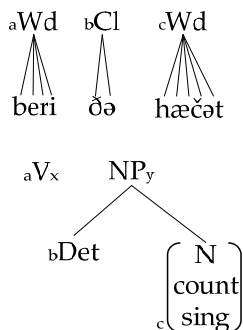
The mapping which links the referential tier and tiers of the CS could be regarded as a part of the conceptual integrative mechanism, which integrates the context with the mental information derived from the linguistic expression itself. Jackendoff (personal communication) describes this integration as adding descriptive features of the referent to the descriptive features of the token which corresponds to this referent

in CS. The former descriptive features can be derived via perception, or from one's f-knowledge, which, according to Jackendoff, contains one's sense of the communicative context, including one's sense of one's interlocutor's intentions (Jackendoff 2002b: 273) and possibly one's awareness of the referent, which for the present theory is a particularly important part of context.

### **3.3 Treatment of idioms and constructions within the framework of Conceptual Semantics**

Jackendoff (1995: 136; 1997: 157; 2002b: 167, 178) suggests that idioms, constructions and other fixed expressions larger than words are listed in the lexicon. He appeals to Representational Modularity, according to which anything that links phonology, syntax and meaning is the responsibility of the correspondence rule modules and therefore a part of language (Jackendoff 1997: 157). "There is no other faculty of the mind in which they can be located" (ibid. 158). He also points out that in order to draw a boundary between the lexical theory and the theory of fixed expressions, one must show how the latter is distinctively different from the former (ibid. 157). Jackendoff (2002b: 153) operates with the term lexical item, which is quite different from the notion of word and denotes any item made up of linguistic parts and stored in long-term memory. He argues (ibid. 65, 154) that fixed expressions and idioms have to be stored in long-term memory as units, since their meaning cannot be predicted from the meaning of their parts. It seems that for Jackendoff (ibid. 186) there are no clear boundaries between syntax, morphology and lexicon as he states that e.g. a syntactic phrase-structure rule is a lexical item. Jackendoff (ibid. 154-162) also discusses lexical items smaller than words, i.e. stems and affixes.

Jackendoff (2002b: 170) proposes treating idioms as lexically stored phrasal units with bifurcated mapping: in his formal descriptions, which include an idiom's phonological structure, syntactic structure and LCS, phonology-syntax connections are denoted by pre-subscripts and syntax-semantics connections by post-subscripts. Idioms like those presented in Figure 44 and Figure 45 below have a regular phonology-syntax mapping but an irregular syntax-semantics mapping: the words do not contribute individually to the meaning.



[RECONCILE ([ ]<sub>A</sub>), [DISAGREEMENT]<sub>y</sub>]<sub>x</sub>

Figure 44 Jackendoff's (1997: 168) description of the English idiom *BURY THE HATCHET* 'reconcile a disagreement'

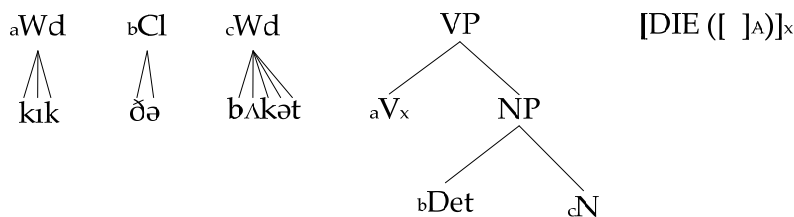


Figure 45 Jackendoff's (1997: 169) description of the English idiom *KICK THE BUCKET* 'die'

Jackendoff (1997: 168-169) attempts to capture the syntactic mobility of *BURY THE HATCHET* 'reconcile a disagreement' as opposed to the certain fixedness of *KICK THE BUCKET* 'die'<sup>97</sup>. *BURY THE HATCHET* (Figure 44) has some metaphoric semantic composition and its LCS can be partitioned into chunks that correspond to subidiomatic readings of constituents: *bury* means 'reconcile' and *the hatchet* means 'disagreement'. The subscript *x* on the whole LCS maps it into the verb, while the second argument of LCS is mapped into an NP by the subscript *y*. The syntactic structure does not stipulate a VP constituent: the V and the NP are not syntactically connected in any way. As a result *the hatchet* is movable. On the other hand, in *KICK THE BUCKET* (Figure 45), *bucket* has no independent meaning and no  $\theta$ -role, therefore it has to be linked to *kick* syntactically.

<sup>97</sup> Section 4.2 presents some evidence of this PU passivization, which demonstrates that even semantically non-decomposable PUs allow some transformational productivity. This evidence suggests that syntactic variability and transformational deficiencies of non-compositional PUs cannot be predicted solely on the basis of their semantics.

Unlike Jackendoff (1997, 2002b), Nikanne (2005a) assumes that lexicon, morphology and constructions belong to separate modules. According to Nikanne (2005a), phonological, syntactic, conceptual and other representations are products of representational modules, which define their well-formedness. Morphology, lexicon and constructions are mapping modules, which do not include any representations of their own, but instead specify how to map phonological, syntactic and conceptual structure representations onto each other (Nikanne 2005a: 192-197). In Nikanne (2005a: 199), the mapping module Constructions encompasses linking devices that license irregular (construction-specific) syntactico-semantic mapping. Since constructions are higher level items built of lower level lexical and morphological items, there is good reason to treat them as distinct systems (ibid. 198, 210). Even though both lexical items and constructions are structure-specific linking devices, they differ in the sense that lexical linking between syntactic and conceptual structure is less fixed and more regular than that of constructional linking<sup>98</sup>. The regular linking principles do not refer to any particular combination of syntactic or semantic categories, particular lexical items or morphological forms (ibid. 210). In Nikanne's (ibid. 199) approach, constructions are linking devices that license irregular linking patterns.

In this respect, treatment of constructions in Nikanne's (2005a) version of Conceptual Semantics is similar to that of Goldberg (1995), who presents the following definition of a construction:

C is a CONSTRUCTION iff<sub>def</sub> C is a form-meaning pair  $\langle F_i, S_i \rangle$  such that some aspect of  $F_i$  or some aspect of  $S_i$  is not strictly predictable from C's component parts or from other previously established constructions (Goldberg 1995: 4).

However, Nikanne (2005a: 198) points out that the problem with this definition is that it does not clarify how the meaning of a complex syntactic form could be predictable from its component parts in those cases when the pair  $\langle F_i, S_i \rangle$  is not a construction. The major difference between Conceptual Semantics and constructional approaches is that in Conceptual Semantics, all linking is not assumed to be governed by constructions. Thus, Conceptual Semantics strives to keep irregular constructions apart from regular syntactico-semantic linking patterns.

---

<sup>98</sup> For a more detailed discussion on regular linking in lexical items see e.g. Nikanne (1997, 1998 and 2005).

I assume that the notion of constructions (in the sense of Nikanne) and of PUs basically overlaps. PUs also license irregular (idiosyncratic) linking between different modules and levels of representation: lexical and morphological items (which are themselves mappings between semantic, syntactic and phonological levels of representation), syntactic and conceptual configurations (also decomposable to further levels). Irregularity can be understood as case-specific exceptions to the principle of the generality of linking rules governing grammatical and semantic structures. At least some aspects of a PU (and its behaviour) cannot be predicted by the general rules of the syntactico-semantic linking.

Following Nikanne's (2005a: 193) monostratal theory of representations, I assume that there is only one level of conceptual representation. Insofar as meaning in Conceptual Semantics is equalled to conceptualization, one cannot assume that different meanings, i.e. different conceptual systems, exist to encode lexical and phraseological items, as has been stipulated in the theory of phraseology (see Section 2.1.1). Therefore, the controversial question of the difference of lexical and phraseological meanings, which for decades has occupied the minds of scholars working in the theory of phraseology, is irrelevant within the scope of the theory of Conceptual Semantics. Instead I believe that the difference is in the character of linking. In order to demonstrate this difference, one will have to introduce a new method of formal analysis of PUs, which will include all relevant levels of representation: syntactic, phonological, morphological, conceptual etc. One also has to elaborate on the linking system itself. In the following sections the Tiernet model will be applied to the needs of PU analysis.

### **3.4 The network structure of PUs**

PUs specify linking between many different levels of representation: default phonological and morphological form of lexical constituents, syntactic and conceptual configuration, social and cultural knowledge, subjectivity etc. Applying the Tiernet model to the formal description of PUs' internal structure will result in a complicated network consisting of interconnected tiers. Each tier is an autonomous micro-module with its own primitives and formation rules (Nikanne 2008a, 2008b). Tiers that organize the conceptual structure and link it to syntax have already been mentioned in Section 3.2 (the theta-argument tier, the f-chain, the action

tier, the temporal tier, the semantic field tier, the modal tier, the DA-tier, the syntactic argument tier, the referential tier). In what follows, I will present an overview of some other tiers relevant for PU analysis that together constitute a PU's network structure: the phonological and the morphological tiers (Sections 3.4.1.1 and 3.4.1.2) and tiers of lexical conceptual structure of nouns (Section 3.4.1.3).

The problem of literal meaning and PU analyzability will be discussed in Sections 3.4.3.1 and 3.4.3.3. It will be argued that the formal description of PUs has to include two parallel conceptual structures – CS licensed by the regular linking rules and the CS/PU licensed by the PU itself (Section 3.4.3.2), as well as two parallel referential structures (Section 3.4.3.4). I will present my view of analyzability as PU referentiality of syntactic constituents.

### 3.4.1 Lexical constituents and their description

Following Nikanne (2002; 2005a: 195f), I assume that lexicon and morphology do not have representations of their own; they are linking mechanisms which specify how particular fragments of different tiers are to be linked in a given language. Lexical items interface information between both linguistic and non-linguistic representations (Jackendoff 1997, Nikanne 2000). The only strictly linguistic information is encoded in phonology, syntax and conceptual structure (CS). Phonological structure is needed for pronunciation of the word and identification of its boundaries, syntactic structure determines the word's syntactic behaviour and the CS enables the concept to function either as a predicate or as an argument of some predicate (Nikanne 2000: 80).

Figure 46 and Figure 47 below is a formal description of the three separate word forms: the Finnish verb *heittää* 'throw<sub>WPRES ACT SG3</sub>' and two Finnish nouns *helmiä* 'pearls<sub>SPL PTV</sub>' and *sioille* 'pigs<sub>SPL ALL</sub>'. It includes tiers encoding phonological, morphological, syntactic and conceptual information. In the following subsections I will take a closer look at different parts of this description and linking between them.

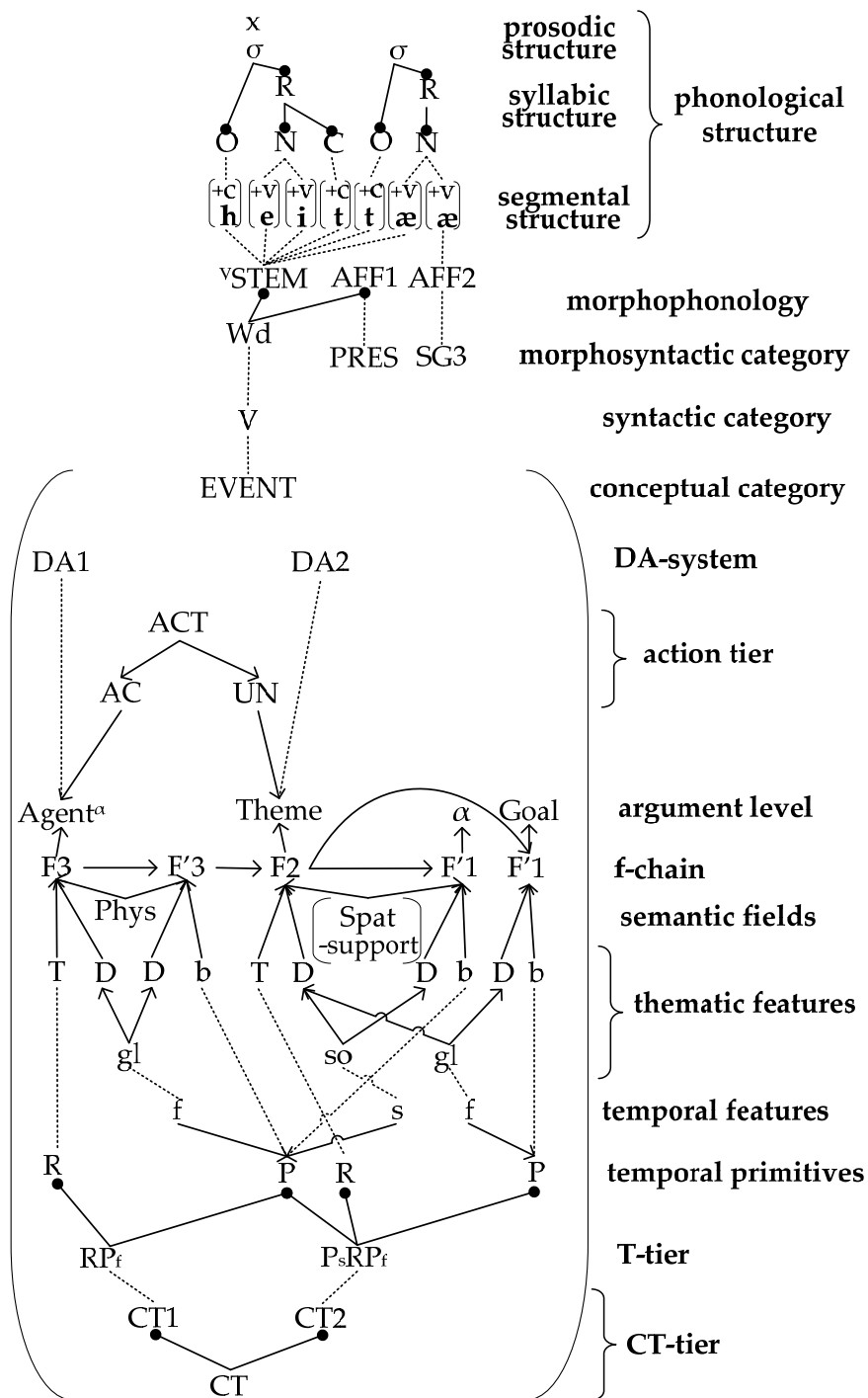


Figure 46 Word form *HEITTÄÄ* 'throw<sub>PRES-ACT-SG3</sub>' and its linking to LCS of the lexical item *HEITTÄÄ* 'throw'

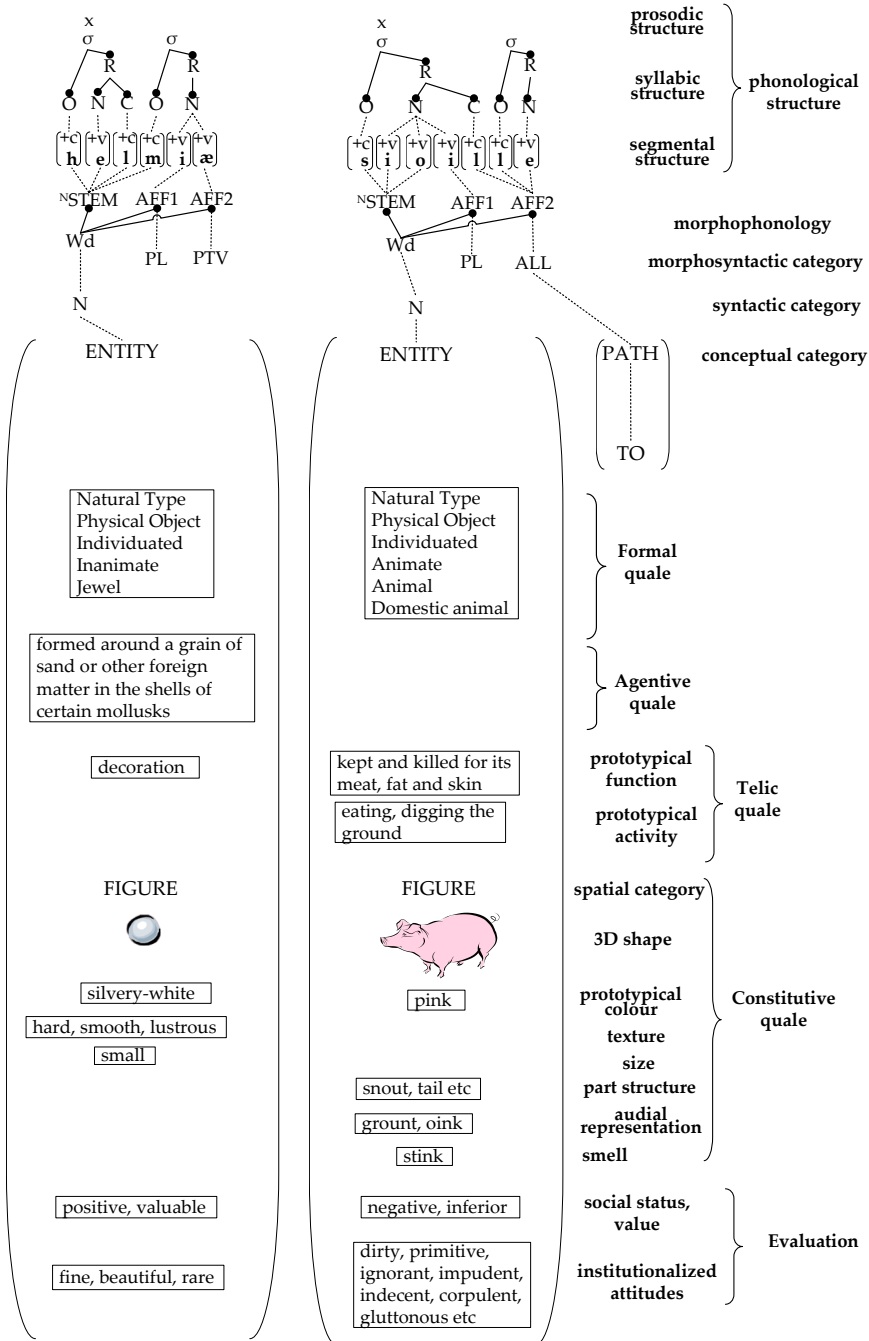


Figure 47 Word forms *HELMIÄ* 'pearl<sub>PL-PTV</sub>' and *SIOILLE* 'pig<sub>PL-ALL</sub>' and their linking to LCS of lexical items *HELMI* 'pearl' and *SIKA* 'pig'



### 3.4.1.1 Tiers of phonological form

At the top of Figure 46 and Figure 47 there is a stripped down representation of the phonological structure (PS) associated with these words. It consists of segmental structure, syllabic structure and prosodic structure, each divided into different tiers. The segmental structure includes a string of sounds (only one distinctive feature is included here – either consonantal (+c) or vocalic (+v); for a more complete set of distinctive features in segmental structure, see Jackendoff 2002b: 7). Syllabic structure shows how individual sounds are combined into syllables ( $\sigma$ ), as well as hierarchical distinctions inside the syllable, i.e. Nucleus, Coda, Onset and Rime parts. Nucleus (indicated by N) – the sonorous core around which the syllable is built (usually a vowel, but also sonorant consonants), and Coda (C) – any material following the Nucleus, are grouped as the Rime (R) – the part of the syllable that remains the same in rhymes. The material before the nucleus is grouped as the Onset (O) – the part that remains the same in alliteration. Above the syllabic structure there is a tier of prosodic structure, which consists of intonational phrases, indicating pauses in pronouncing the syllables, and the metrical grid (indicated by  $\times s$ ), which encodes the relative stress of syllables (Jackendoff 2002b: 6-8).

### 3.4.1.2 Tiers of morphophonology and morphosyntax

According to the Correspondence Rule Strategy (Jackendoff 1990: 156-157), lexical items do not completely specify how different tiers are to be linked to each other; this is a priority of extra-lexical correspondence rules. One of these rules – the principle of default linking of DAs to syntactic arguments – has already been mentioned in Section 3.2.8 above. Another set of correspondence principles can be detected e.g. in the inflectional morphology of nouns in Finnish and is formulated as follows:

- The constituents of an inflectional form of a word (Wd), which belongs to a syntactic category noun (N), are the noun stem (<sup>N</sup>STEM) and inflectional affixes<sup>99</sup> (AFF) which represent inflectional categories.
- The inflectional categories of nouns are number (NUM), case (CASE) and possessor (POS).

---

<sup>99</sup> This term is used here mostly in the same sense as the Finnish term *taivutustunnus* is used in Hakulinen et al. (2004: 85).

- One inflectional form of a word can contain only one AFF of the same category (e.g. there cannot be several case AFFs). Thus the maximal number of AFFs of a noun in Finnish is 3.
- Potential AFF positions are marked from left to right as AFF1, AFF2 and AFF3. AFF1 is the closest position to the stem; AFF3 is the most distant one.
- Potential inflectional categories are ordered according to the implicational hierarchy NUM > CASE > POSS. The highest category in the hierarchy gets the closest position to the stem.

Similar principles can be established for the inflection of Finnish finite verbal forms. The potential inflectional categories here are passive (PASS), tense, mood (T) and person (PERS), which form an implicational hierarchy PASS > T > PERS. Linking principles for Finnish nouns are illustrated in Figure 48 and for finite verbal forms in Figure 49 below.

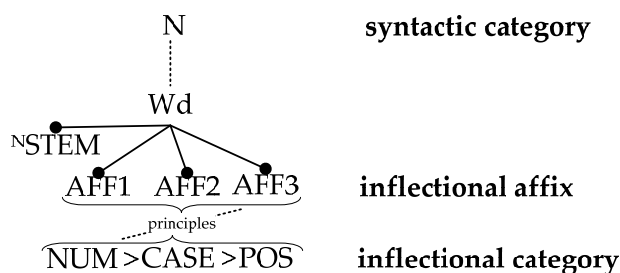


Figure 48 Principles of inflectional affixation in Finnish nouns

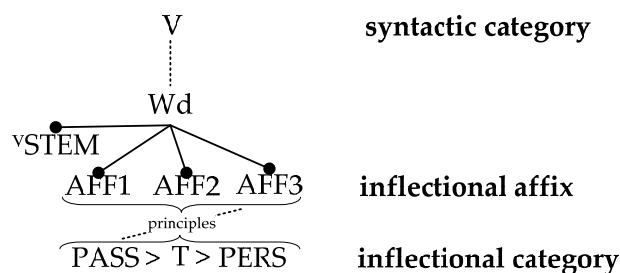


Figure 49 Principles of inflectional affixation in Finnish finite verbal forms

Going back to Figure 46 and Figure 47, the phonological string of the segmental structure interfaces with the morphophonological structure, which combines sounds into word forms. These word forms are in turn constituted by a stem and inflectional affixes. The latter morphophonological constituents are mapped onto morphosyntactic categories: an affix corresponds to some inflectional category according to

the principles described above. A morphophonological unit as a whole is connected to syntax via a correspondence link to some syntactic category such as Noun or Verb.

### 3.4.1.3 Tiers of the noun's Lexical Conceptual Structure (LCS)

Below the level of morphosyntactic categories in Figure 46 and Figure 47, there are syntactic categories that interface with conceptual structure by correspondence to some conceptual category: the verb corresponds to the conceptual category SITUATION and Nouns correspond to the conceptual category ENTITY. The allative case corresponds to the category PATH and the thematic function TO on the f-chain level. The internal structure of the category SITUATION is presented according to Nikanne (1990b, 1995, 1997a, 2005a).

Both Nikanne and Jackendoff argue that non-linguistic fragments of representation of the word (audio-visual, spatial, haptic, social etc.) can be regarded as parts of its meaning. Jackendoff (2002b: 347f) claims that CS and spatial structure (SpS) together compose a concept or meaning; these two levels interface with each other and together they conceptualize the world. CS conceptualizes predicate-argument relations, category membership, type-token distinction, quantification etc. SpS encodes the spatial understanding of the physical world, supports visual and sensory object categorization and identification (ibid. 346f): the entity's dimensionality, shape, size, colour, texture, weight, smell, part structure etc. Some lexical items lack the SpS component, e.g. abstract and logical concepts have only CS. On the other hand, many perceptual properties, e.g. those denoting colours, can be encoded directly in SpS. Jackendoff (ibid. 350) remarks that CS and SpS overlap via notions of physical object, part-whole relationships, locations, force and causation, which appear in both systems. Since there is a great deal of interaction between these two components, Jackendoff (ibid.) is naturally concerned with the question of their boundaries. He asks: "how far can content be bled out of CS into SpS?" This problem seems to be irrelevant in the Tiernet model, since we are not dealing with large modules anymore, but instead with micro-modular tiers, each of which conceptualizes only a particular level of representation. Lexical items establish connections only between tiers which are relevant for conceptualization of a particular information structure.

### 3.4.1.3.1 Qualia structure

The decompositional semantics of a nominal (as is done with the nouns *helmi* ‘pearl’ and *sika* ‘pig’ in Figure 47) is divided into several tiers, which are grouped according to Pustejovsky’s (1995, 2001, 2003) qualia structure. Pustejovsky introduces four qualia (Jackendoff 2002 does not take this to be an exhaustive list), which are different types of properties of lexical concepts and are based on classic Aristotelian modes of explanation for an entity or relation:

- Formal quale distinguishes the object within a larger domain, including its taxonomic structure.
- Agentive quale encodes information about how an entity comes into existence, factors involved in the origin or “bringing about” of an entity, information about its life-cycle stages. Jackendoff (2002: 370) points out that this quale might also encode information about what an entity will develop into.
- Telic quale specifies information about activities in which the entity takes part, its purpose and function.
- Constitutive quale includes information about the objects structural attributes: dimensionality, shape, size, colour, texture, weight, smell, the material it is composed of, its part structure, etc.

Pustejovsky (1995: 428) argues that such structured information is not only useful for nouns, but also necessary to account for their semantic behaviour.

There are some differences in Jackendoff’s (2002b) and Pustejovsky’s (1995) treatment of qualia. Pustejovsky (1995: 427) places information about an object’s shape, dimensionality, colour, etc. in the Formal quale, while Jackendoff (2002b: 370) relates such structural attributes to the Constitutive quale. Jackendoff (ibid. 370) also remarks that much of the information in the Constitutive quale will interface with the Spatial Structure. According to Jackendoff (ibid. 369), the Formal quale includes the taxonomic structure, for instance *pig* is a material entity, animate, a kind of animal etc. For discussion on taxonomic structure and problems associated with it see e.g. Jackendoff (ibid. 343-345). In my analysis I will mostly follow Jackendoff’s version.

### 3.4.1.3.2 Evaluation

The CS of nouns in Figure 47 also includes a tier labelled as Evaluation, which encodes not the perceptual qualities of an entity, but the

associated, attitudinal ones, e.g. the social status and value of the entity and various institutionalized qualities and properties, attributed to it. This category differs from Jackendoff's (2002b: 312-313) valuation features in a cognitive structure of a percept, which include such distinctive pairs as external vs. internal, familiar vs. novel, self-produced vs. non-self-produced, meaningful vs. non-meaningful and mattering vs. non-mattering. Instead, it can be compared to the notion of connotation<sup>100</sup> as defined by Apres'jan (1974: 159), i.e. "stable features of the concept, which embody the evaluation of a respective object or a fact of reality adopted in a given language community". There is, however, another theoretical framework which is concerned with evaluation and which can be helpful in my analysis – the Appraisal framework (Martin 2000; Martin & Rose 2003; White 2002, 2005). Appraisal is a very broad term, which encompasses all evaluative uses of language, attitudinal positioning being one of its basic options. The three sub-types of Attitude are (White 2002, 2005):

- Affect – the writer/speaker is indicating his/her emotional disposition towards the ENTITY or SITUATION.
- Judgment – normative assessments of human behaviour or character typically making reference to some system of social norms or conventions.
- Appreciation – assessments of semiotic and natural phenomena by reference to their value in a given field, most typically by reference to their aesthetic qualities.

All three can be positive (+) or negative (-). In Figure 50 below, the lexicalized evaluative properties that are present in the LCS of *helmi* 'pearl' and *sika* 'pig' are grouped according to these sub-types. The evaluative properties of *helmi* 'pearl' fall into the scope of positive appreciation, while the evaluative properties of *sika* 'pig' are distributed between negative Appreciation and negative Judgment.

---

<sup>100</sup> See Section 2.1.3 in Chapter 2 of this book for a broader discussion of the notion of connotation in linguistic literature.

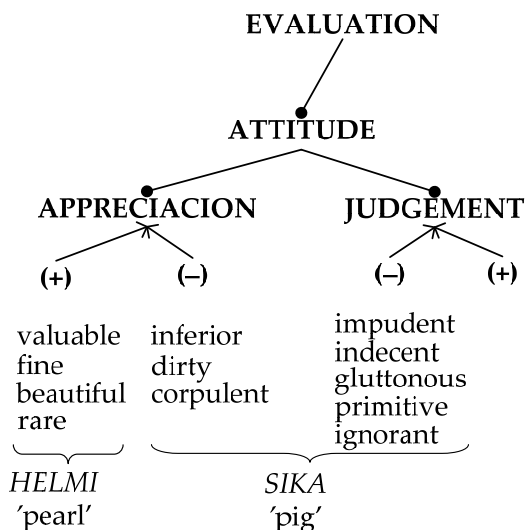


Figure 50 Evaluative aspects of the concepts HELMI 'pearl' and SIKI 'pig' presented as positive and negative Appreciations and Judgments

Despite the fact that the behaviour normatively assessed in Judgment is human, it is possible to apply this notion to the evaluative properties of *sika* 'pig', since we are dealing with an application to an anthropomorphized animal of the same value and normative system which humans apply to each other. Some of the institutionalized attitudes towards pigs are based on fragments of encyclopedic knowledge about these animals: e.g. gluttony and insatiability are attributed to pigs based on the knowledge of their omnivorous diet, and attributed dirtiness comes from their habit of wallowing in mud. Ironically, this latter property has been extended to metaphorically denote the phenomenon of human sweating (e.g. Finnish language has a conventional simile *HIKOILLA KUIN SIKI* 'sweat like a pig' to denote sweating profusely), which belongs to our cultural concept of dirt (Lagerspetz 2006) but contradicts the zoological knowledge about pigs, who have ineffective sweat glands and use mud only in order to lower their body temperature. It seems that natural experience as a source of our Evaluation of pigs is to a large extent interwoven with fragments of symbolic knowledge coming from culturally relevant sign systems (mythology, tales, fables, religious texts) (Dobrovol'skij & Piirainen 2005: 96). Symbolic associations as well as the social status of pigs may be different in different cultures.

Pigs are in fact symbolically ambiguous: the positive attitude towards them in pagan mythology (in Sumerian-Semitic cults the pig is an

attribute of mother-goddesses and in Greek mythology it was a sacred animal of the fertility goddess Demeter; also in Oceanic, Celtic and Scandinavian cultures it was a symbol of fertility) contrasts with basically negative symbolism in world religions (in Judaism and Islam pigs are unclean, in Buddhism they symbolize ignorance and in Christianity the pig is the symbol of the devil, temptation, voracity, insatiability and ignorance). The association with ignorance is entirely symbol-based, since in reality pigs are intelligent species with a cognitive capacity that ranks them as the fourth smartest animal group in the world. They are definitely smarter than owls; nevertheless the latter are conceptualized in Western culture as a symbol of wisdom and knowledge due to their association with the Greek goddess Pallas Athena. A manifestation of the symbolical association of pigs with ignorance can be found in literature e.g. in the fable СВИНЬЯ ПОД ДУБОМ [*Svin'ja pod dubom*] 'The Pig under the Oak Tree' (1821-1823) by the 19<sup>th</sup> century Russian writer of fables in verse Ivan Krylov. The fable tells a story of a pig rooting the earth under an oak tree. Ignoring the fact that this could damage the tree's roots, it said that it did not care about the tree itself, but was interested only in acorns. If the pig could look higher than its nose, it should have seen that it was the oak tree that delivers the acorns. Krylov concludes his fable with a moral: an ignoramus criticizes sciences, education and knowledge without understanding that he reaps their fruit.

### 3.4.1.3.3 Motivational correspondence links between a PU's CS and the CS of its lexical constituents

The evaluative tier in the semantic description of lexical items, noun constituents of the PU, can provide a source of motivational correspondence links between the LCS of a constituent and the CS licensed by the PU (CS/PU). In the case of the Finnish PU X HEITTÄÄ HELMIÄ SIOILLE |  $NP_{SUBJ}[N\{X\}]$   $V\{heittää\}$  'throw'  $NP_{OBJ}[N_{PL}\{helmi\}$  'pearl']  $PP_{ALL}[NP_{PL}\{sika\}$  'pig'] | lit. 'X throws pearls to pigs', id. 'X causes a transfer of some entity, evaluated by the speaker as good, to a recipient, evaluated as inadequate in some way', the evaluative properties that we have observed in conceptual structures of the lexical items *helmi* 'pearl' and *sika* 'pig' motivate properties of the thematic arguments which appear in the CS/PU and which are co-referential<sup>101</sup> with the syntactic constituents

<sup>101</sup> See Section 3.4.3.4 of this chapter for an account on referentiality of PU constituents.

(nouns) that correspond to the phonological and conceptual structures of these items. Thus, the evaluative properties of *helmi* 'pearl' (positive, valuable, fine, rare) motivate the positive appreciation GOOD in the first of the two characterizing propositions (Figure 51 below), while the evaluative properties of *sika* 'pig' (negative, inferior, primitive, ignorant) motivate the negative judgment INADEQUATE in the second proposition.

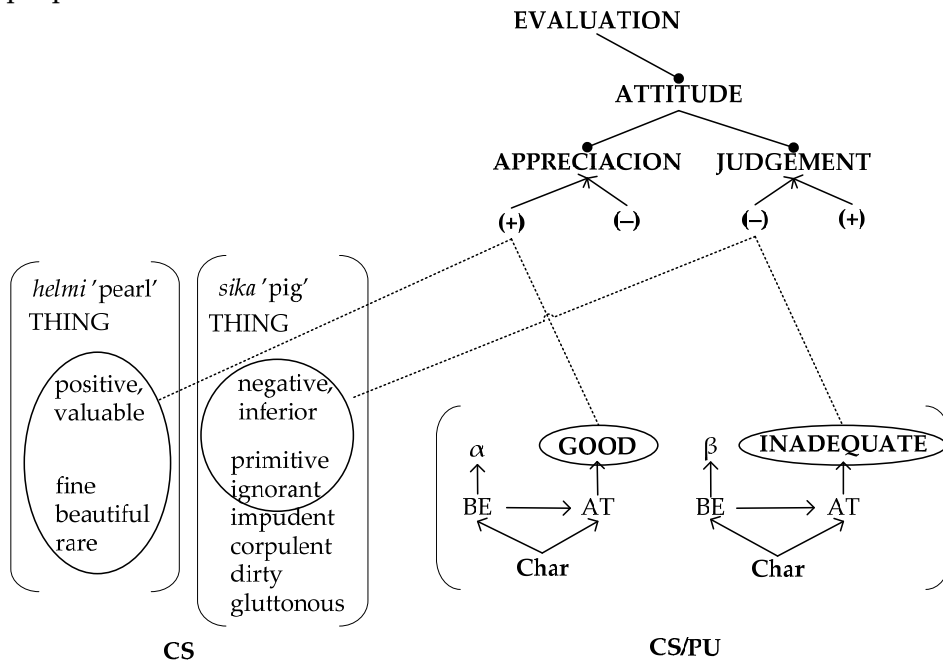


Figure 51 Motivational correspondence links between the parts of the CS/PU and the lexical components' CS in the Finnish PU HEITTÄÄ HELMIÄ SIOILLE 'throw pearls to pigs'

The same evaluative properties primitive, ignorant are present e.g. in the following Russian (65), (66) and Finnish (67) PUs:

- (65) ПО ОБРАЗУ – КАК Я, А ПО УМУ – СВИНЬЯ (Dal' 1994)  
 [Po óbrazu – kak ja, a po umú – svin'já]  
 lit. 'An image like mine, but an intellect like a pig's'  
 id. 'a primitive, ignorant person'
- (66) ПРОСТ, КАК СВИНЬЯ, А ЛУКАВ, КАК ЗМЕЯ (Dal' 1994)  
 [Prost kak svin'já, a lukáv kak zmejá]  
 lit. 'Simple as a pig, but cunning as a snake'  
 id. 'a primitive but cunning person'



(67) *VIE SIKI SAKSAAN, TUO SIKI SAKSASTA - SIKI ON SIKI JA PANNOO VUAN: "NOH, NOH"*<sup>102</sup>

lit. 'Take a pig to Germany, bring a pig from Germany – a pig is a pig and makes only "oink, oink"'

id. 'One cannot make a stupid person into something else, not even by sending him abroad'

It is notable that only a part of the properties commonly associated with a *pig* is involved in motivational correspondences: thus, properties like obesity, dirtiness, coarse indecent behaviour and gluttony are not mapped onto this particular CS/PU; however, they can be activated in other PUs. For example, the Russian PU ПОСАДИ СВИНЬЮ ЗА СТОЛ, ОНА И НОГИ НА СТОЛ [*posadí svin'jú za stol, oná i nógi na stol*] lit. 'set a pig to sit at the dinner table, it (will put) its legs on the table', 'id. an impudent person who has bad manners, behaves in a rudely familiar way' activates a correspondence between the negative judgment IMPUDENT – a property of the thematic argument co-referential with the lexical component *свинья* [*svín'ja*] 'pig' in the CS/PU and the institutionalized associations impudent, coarse in the CS of this lexical unit (Figure 52 below).

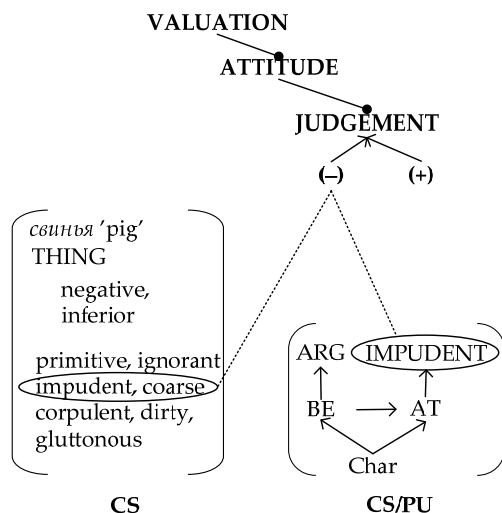


Figure 52 Motivational correspondence links between the part of the CS/PU and the lexical component's CS in the Russian PU ПОСАДИ СВИНЬЮ ЗА СТОЛ, ОНА И НОГИ НА СТОЛ [*posadí svin'jú za stol, oná i nógi na stol*] lit. 'set a pig to sit at the dinner table, it (will put) its legs on the table too', id. 'an impudent person who has bad manners, behaves in a rudely familiar way'

<sup>102</sup> Sananparsikokoelma (2010) - Riistavesi, A. Ihalainen, 1932

The concepts of dirtiness and indecency, which are both associated with pigs, are metaphorically connected with each other. Like dirt which exists in relation to the idea of what the objects should be like (Lagerspetz 2006), indecency – the metaphorical moral dirt – is deviance from generally accepted behaviour: we speak of obscene or indecent movies and jokes as *DIRTY MOVIES* and *DIRTY JOKES*, a malicious or scandalous lie is a *DIRTY LIE*, unethical or corrupt politics are *DIRTY POLITICS*, money acquired by illicit or improper means is *DIRTY MONEY*, etc. This connection is expressed in PUs which contain components *pig* and *dirt (dung)* and denote deviant behaviour of a human, e.g. the Russian proverbs НАПЯДИ СВИНЬЮ В СЕРЬГИ, А ОНА В НАВОЗ [*nar'jadí svin'jú v sér'gi, a oná v navóz*] lit. 'adorn a pig with earrings and it (will go) into dung', id. 'no matter how you try to change someone, his true nature and habits will surface', and СВИНЬЯ (ВЕЗДЕ/ВСЕГДА) ГРЯЗЬ НАЙДЕТ [*svin'já vezdé/vs'egdá gr'az' najd'ót*] lit. 'a pig will always find dirt', id. 1. 'someone who likes to do disgraceful things will always find a chance to do them'; 2. 'a person with some negative qualities will see the same qualities everywhere else'.

### 3.4.2 Linking values and the default form of a PU

Each linking type discussed in Section 3.2.1 is able to take different values, which indicate the strength of a particular connection. A value can be either licensed idiosyncratically (by a construction or a construction family), or as a general language rule. The values which I use in the current model are fixed (Fxd), default (Dfl) and non-default (NDfl) (Figure 53 below). A link with no given value indicates that this particular connection is not licensed as default or fixed by any construction, construction family or general grammatical principle, it has no preferred case and therefore can vary freely within the limits and constraints of general mapping principles, which exist between the involved levels of representation. Since it is important to distinguish the fixed, default or non-default linking stipulated by the general rules of grammar of a natural language from such linking licensed by a particular PU of this language, I am going to need two linking value systems: a language-specific one and a PU-specific one. In my formal descriptions the latter will be marked by adding a PU abbreviation in parentheses after a value marker.

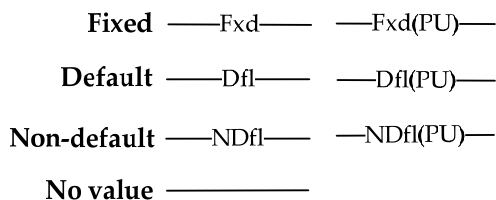


Figure 53 Linking values

A fixed value is the strongest. It indicates that a particular mapping cannot be subject to change and is always the case. As long as there is at least some possibility of variation, the link cannot be regarded as fixed and thus receives a default value<sup>103</sup>, which means that this linking is the preferred case, but it can be violated i.e. become subject to variation. In order to distinguish between preferred and less common realizations of any particular mapping, the latter will be marked as non-default. Fixedness as a value cannot be relative: by postulating a fixed link, I assume that this particular mapping is not subject to variation. The opposite of complete fixedness is total non-acceptability of a link. A possible way to model the dimension of strength of connection is to represent it as a continuum which stretches between completely unacceptable “0” and completely fixed “1” linking with non-default, free and default values in the middle (Figure 54 below).

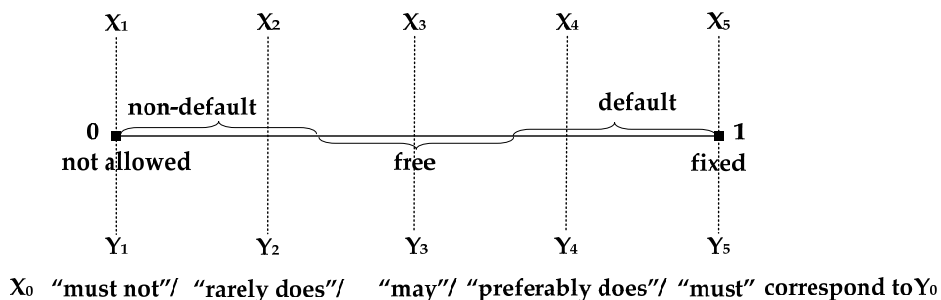


Figure 54 A continuum of connection strength

Thus, relativity in strength can be hypothesized for the default value in a sense that, as long as a particular mapping is perceived as being more preferable than any other possible mapping between two given levels of

<sup>103</sup> Culicover and Jackendoff (2005: 164) formalize a default correspondence between CS and SS as  $CS \Leftrightarrow_{\text{default}} SS$ , where  $\Leftrightarrow$  indicates that correspondence is licensed and the subscript *default* indicates that this correspondence is the preferred or unmarked case but can be violated.

representation, this preferability can be stronger or weaker. The strength of links can be measured by e.g. turning to corpus statistics, and comparing the frequency of occurrences, which different realizations of any particular mapping have in relation to each other. As long as no statistically significant difference between frequencies of variants of the same mapping can be detected, the link can be considered free. The picture in Figure 54 should be treated as a tentative explanatory model and does not provide any exact parameters that would allow to establish border values between e.g. non-default and free, on the one hand, and free and default, on the other. Such values must be determined by means of statistical significance analysis, which is well beyond the scope of the current study.

By introducing a distinction between different linking values and assuming relative strength for some of them, one gains an efficient tool for PU variation analysis: strength can be expected to vary between different language speakers and different registers, as well as to become subject to diachronic changes, which can either weaken or strengthen the link. Thus, not only can one account for variability in PUs, but also one is able to take into consideration the fact that acceptance of variability can actually differ in different individuals. For instance, if a language speaker perceives all over mappings except one as non-acceptable for a certain PU, his/her value of this mapping is obviously fixed, while other speakers may value this mapping as default and therefore accept some other mappings as well. In both cases we are dealing with intuitive *a priori* judgments. However, in order to arrive at a generalization, the strength of the link has to be verified against corpus data. It is quite common that data obtained from large corpora disproves statements about PU fixedness, which have been based solely on intuition.

Since my model tends towards a truly atomistic representation, values are given to each and every link separately, i.e. instead of labelling an entire word form or an entire syntactic configuration as fixed or default, I am rather looking at a combination of default or fixed links, which make up this word form or syntactic configuration. This allows me to achieve a more flexible and accurate description of a PU's structure. Figure 55 below demonstrates that a PU's interconnected syntactic (SS), morphological (MS) and phonological structures (PS) together form a network, which can feature both language-specific and PU-specific free, default and fixed linking. The Finnish PU examined here is X

HEITTÄÄ HELMIÄ SIOILLE |  $NP_{SUBJ}[N\{X\}] V\{heittää\text{ 'throw'}\} NP_{OBJ}[N_{PL}\{helmi\text{ 'pearl'}\}] PP_{ALL}[NP_{PL}\{sika\text{ 'pig'}\}]$  | lit. ‘X throws pearls to pigs’, id. ‘X causes a transfer of some entity, evaluated by the speaker as good, to a recipient, evaluated as inadequate in some way’. Note that the only CS presented in this picture is the rule-based, regular one, while the PU licensed CS is not yet included into this description. The latter CS will be introduced for the first time in Section 3.4.3.1 and discussed in Section 3.4.3.2.

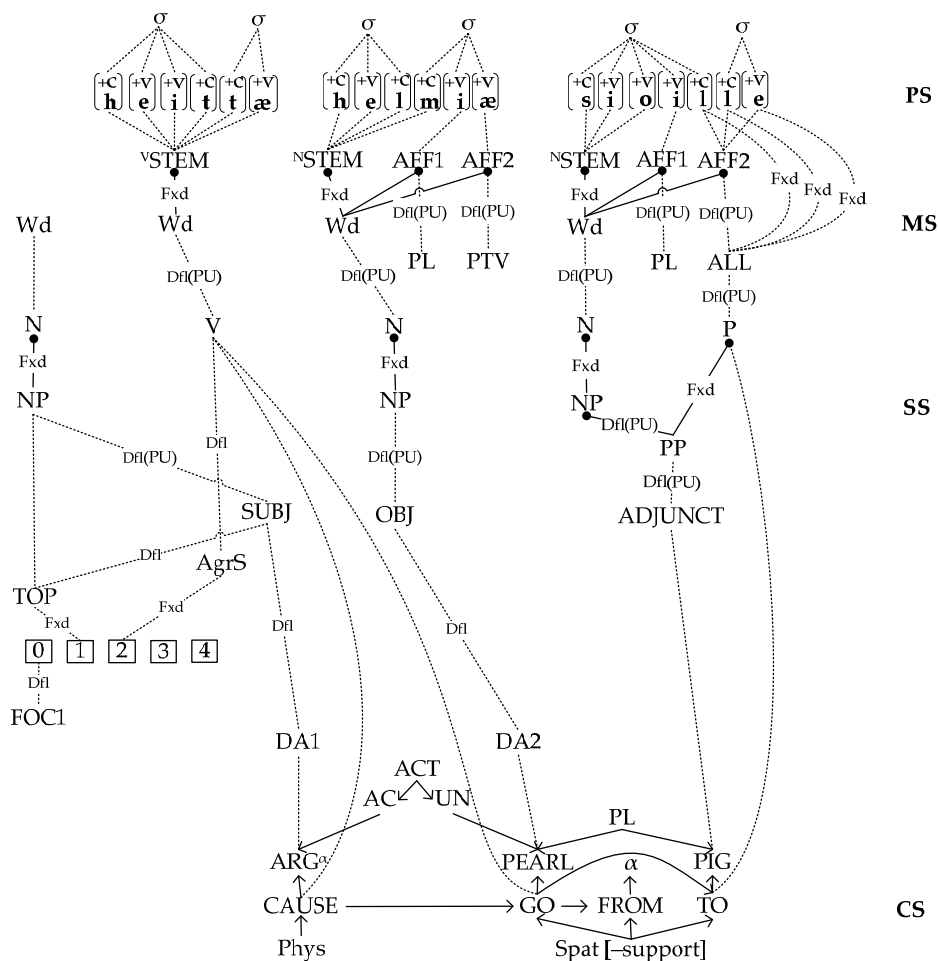


Figure 55 Default phonological, morphological, syntactic and regular conceptual structure of the Finnish PU HEITTÄÄ HELMIÄ SIOILLE ‘throw pearls to pigs’ (the PU licensed CS is not yet included into this description)<sup>104</sup>

<sup>104</sup> Here information structure and its linking to word order tier are formalized according to Nikanne (2008b).

Constituency links between phrases and their heads are marked as fixed, e.g. an NP has to have N as its obligatory constituent; this value is stipulated by the general rules of language and is not licensed by the PU (cf. double vs. single line notation in Jackendoff 2002: 9-10). Another example of language-specific fixed connections is the constituency link drawn between the morphological category Wd and morphophonological category STEM, since a stem is an obligatory constituent of a word. Correspondences between phonemes [l], [l] and [e] in the segmental structure and the allative case in the morphological structure represent an example of language-specific fixed one-to-many mappings.

Since lexical items and their morphology are licensed by the PU, correspondences between syntactic category units (V, N, P) and units of morphological category (Wd, ALL) are marked as Dfl(PU): the verb by default corresponds to the word *heittää* 'throw', and the two nouns to the words *helmi* 'pearl' and *sika* 'pig' respectively, while P corresponds to the allative case. The PU licensed default morphology of lexical items *helmi* 'pearl' and *sika* 'pig' is formalized in such a way that the links between their inflectional affixes AFF1 and AFF2 and the case and number categories are marked as Dfl(PU): in the word form *helmiä* 'pearl<sub>PL PTV</sub>', AFF1 has a Dfl(PU) correspondence with the plural number (PL) and AFF2 has a Dfl(PU) correspondence with the partitive case (PTV)<sup>105</sup>; in the word form *sioille* 'pig<sub>PL ALL</sub>' AFF1 by default corresponds to PL and AFF2 by default corresponds to the allative case (ALL). Despite the fact that the biblical source of this PU *alkää heittääkö helmiänne sikojen eteen* 'do not throw your pearls before pigs' contains the possessively marked word form *helmiänne* 'pearl<sub>PL PTV 2PL</sub>', in modern Finnish the PU does not license for this noun component any default correspondence for the inflectional category POS (possessive), i.e. *helmiä* 'pearl<sub>PL PTV</sub>' can appear with any possessive affix or without it (the latter being the default case).

Not all of the noun constituents have a default counterpart in morphophonology: the correspondence link between the constituent of the subject NP and the Wd node is not licensed by this PU and therefore remains free. The verb is not assumed to have any default inflectional form; therefore its morphophonological structure is represented only by the stem which can be freely inflected. The PU's default syntactic

---

<sup>105</sup> The principles of correspondence between inflectional affixes and inflectional categories in Finnish nouns have been presented earlier in Section 3.4.1.2.

configuration is marked by assigning syntactic roles to its constituents. Thus, the free NP has a Dfl(PU) correspondence to the subject, NP *helmiä* ‘pearl’ to the object and PP *sioille* ‘to pigs’ to the adjunct. The NPs’ further linking to DAs indicates their belonging to the argument structure of the verb *heittää* ‘throw’.

Since the default linking does not exclude a possibility of variation, other morphophonological forms can appear in this structure as well. For instance, in Example (68) the form *sialle* ‘pig<sub>SG ALL</sub>’ features a non-default linking between AFF1 and the SG number (Figure 56 below). Note that the AFF1 that corresponds to SG has no connection to the segmental tier, since it is a null morpheme.

- (68) *Haluatko sinä että minä alan luetella asioita jotka on lainattu Juutalaisuuteen Mesopotamiasta tai Egyptistä? Ei, en viitsi heittää helmiä sialle minäkään.*<sup>106</sup>  
lit. ‘Do you want me to start to list things which have been borrowed to Judaism from Mesopotamia or Egypt? No, I don’t care to **throw pearls to a pig**<sub>SG-ALL</sub> either.’

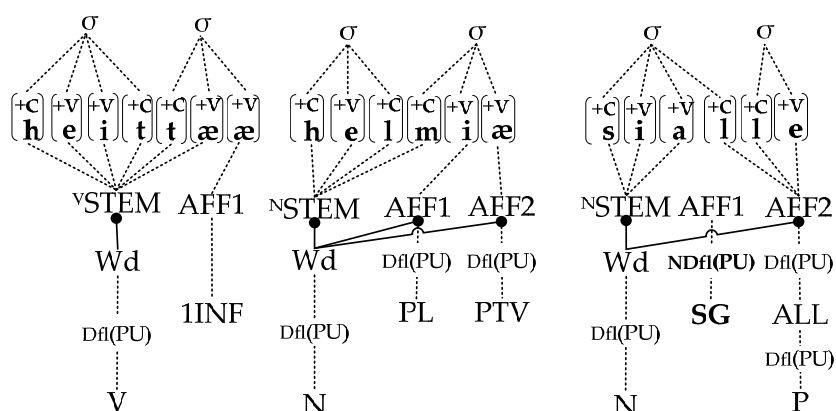


Figure 56 Default lexical item in non-default morphological form (*SIALLE* ‘pig<sub>SG-ALL</sub>’)

In Example (69), two syntactic constituents are mapped to the tier of morphophonology by NDfl(PU) links: instead of the Dfl(PU) linking to *heittää* ‘throw’, V corresponds to *jakaa* ‘distribute’, and N corresponds to *yrittyspomoille* ‘company+boss<sub>SPL-ALL</sub>’ instead of the default *sioille* ‘pig<sub>PL-ALL</sub>’:

- (69) *Ei tarvitsisi jakaa helmiä niille yrittyspomoille ynnämuille, jotka firman piikkiin ostavat lelunsa ja maksattavat puhelunsa.*<sup>107</sup>

<sup>106</sup> GG: sfnet.keskustelu.uskonto.kristinusko/msg/d86b88a778d780c3, Apr 9 1999

<sup>107</sup> GG: sfnet.viestinta.matkapuhelimet/msg/d215b698ce54490d, Sep 21 1998

lit. ‘You wouldn’t need to **distribute pearls to those company bosses and the like**, who at the expense of the company buy their toys and get their phone calls paid for.’

In Figure 57 below these correspondence links are marked with the NDfl(PU) value. Note, that although *yrityspomo* ‘company boss’ is a non-default lexical constituent for this PU, its morphological form is the same as in *sioille* ‘pig<sub>PL-ALL</sub>’ and therefore these particular links retain their Dfl(PU) value. The syntactic structure in the above example also contains a few additional non-default elements: the NP constituent of the PP *yrityspomoille* ‘company+boss<sub>SPL-ALL</sub>’ is modified by a pronoun *niille* ‘those<sub>PL-ALL</sub>’ and the phrase as a whole features a non-default coordinated PP adjunct *ynnä muille* ‘and the like<sub>PL-ALL</sub>’.



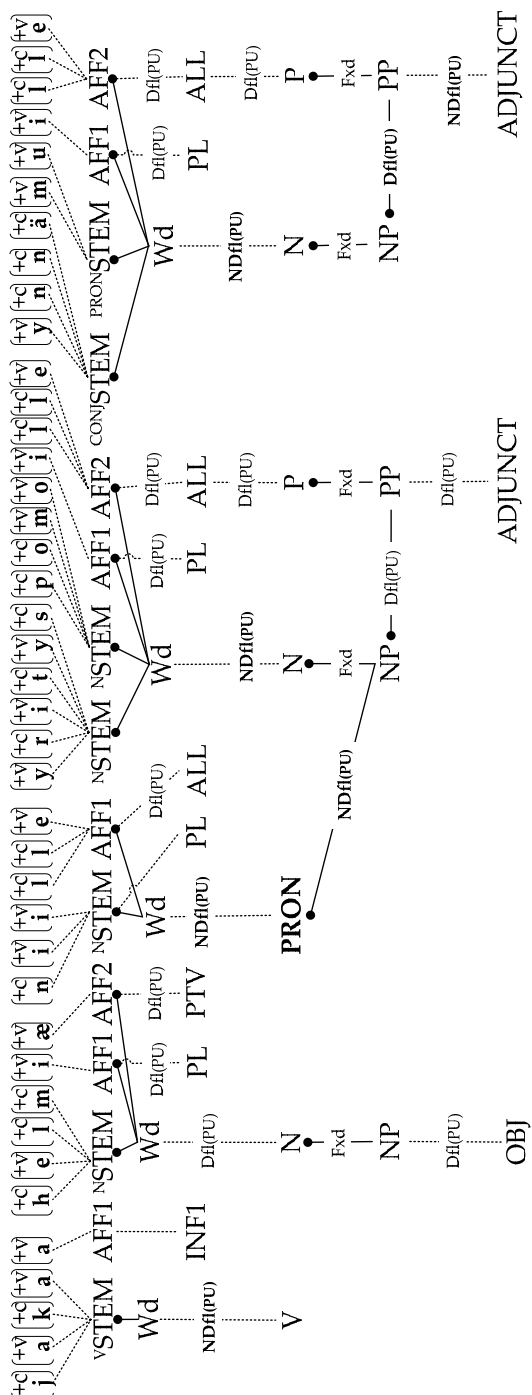
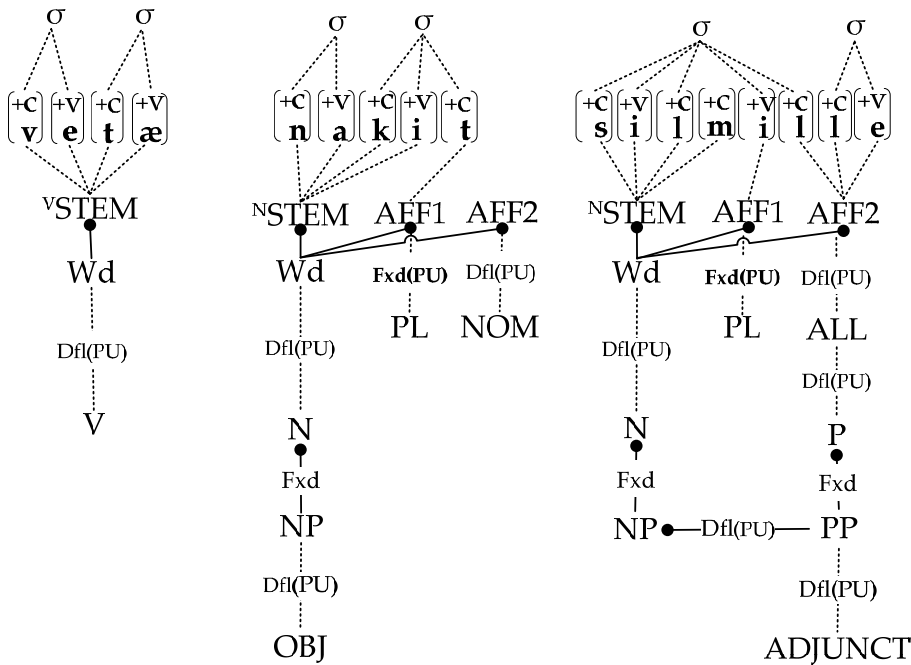


Figure 57 Non-default constituents, default morphology in Example (69)



The nominative case (NOM) is not marked here as fixed for the noun *nakit* 'sausage<sub>PL-NOM</sub>', since the PU is able to appear in negative forms (Example (70) below), where the negated object is marked with the partitive case (PTV):

- In a similar way, the ALL cannot be regarded as a fixed case as long as there is evidence of its variation, e.g. the illative case (ILL) as in Example (71):

<sup>108</sup> <http://kaksplus.fi/keskustelu/plussalaiset/mitas-nyt/1730870-suomi-alkoholistien-paratiisi/sivu5.html#post20482823>, May 14 2010

- (71) *Niin no, tuntuu ainakin hieman vaikealta ruveta asettelemaan nakkeja silmiin kun kurkku on lytyssä ja alkaa silmissä hämärtyä.*<sup>109</sup>  
 lit. 'Well, it feels at least a little bit difficult to begin to **put sausage**<sub>PL-PTV</sub> **eye**<sub>PL-ILL</sub> when your throat is concertinaed and your eyes become hazy.'  
 id. 'Well, it feels at least a little bit difficult to begin to **get drunk** when your throat is concertinaed and your eyes become hazy.'

As concerns syntactic constituents, the *NP<sub>PL</sub>{nakki 'sausage'}* is marked as fixed, and therefore obligatory, while the V and the PP are marked as default, since they can be omitted like in the following Examples (72) and (73):

- (72) *Somalimies pureskelee khatia, musliminainen ympärileikataan, amerikkalaispoika töhrii seiniä ja suomalaisliikemies vetää nakit julkisella paikalla.*<sup>110</sup>  
 lit. 'A Somalian man chews khat, a Muslim woman is circumcised, an American boy smudges walls and a Finnish man **drags sausages** in a public place.'  
 id. 'A Somalian man chews khat, a Muslim woman is circumcised, an American boy draws graffiti on walls and a Finnish man **gets drunk** in a public place.'
- (73) *Vai on Latte vaan ollut ravitsemuksen merkeissä Tallinnassa.. noo ihan hyvä syy. Parempi se, kun että nakit silmille halvalla naapuriviinalla.*<sup>111</sup>  
 lit. 'So Latte has been to Tallinn just for the sake of nutrition.. well, a pretty good reason. Better that than **sausages on eyes** with cheap neighbour booze.'  
 id. 'So Latte has been to Tallinn just for the sake of nutrition.. well, a pretty good reason. Better that than **getting drunk** with our neighbouring country's cheap booze.'

### 3.4.3 PU's conceptual structure and its linking to other levels of representation

#### 3.4.3.1 The problem of literal vs. idiomatic meaning

Idiomatic (or figurative) meaning is usually contrasted with literal meaning. The observation that the intended (or actual) meaning does not coincide with the literal meanings of constituent words is used as one of the definitional criteria for idiomatic expressions (and figurative language in general). The notions of transparency, motivation, semantic

<sup>109</sup> <http://www.potku.net/forum/viewtopic.php?p=204151#p204151>, Mar 11 2007

<sup>110</sup> <http://www.ylioppilaslehti.fi/1995/11/03/muukalaiset-lain-edessa>, Nov 3 1995

<sup>111</sup> <http://forum.stumppi.fi/posts/list/615/203.page#77370>, Oct 1 2008

compositionality or analyzability are also defined via some postulated relations (mapping) between literal meaning and idiomatic meaning. Studies of idiom processing are to a large extent concerned with establishing the difference between comprehension of literal language and comprehension of idioms. Although the notion of literal meaning in its relation to the idiomatic meaning seems to be crucial for understanding many idiom-related phenomena, the concept of literal is itself highly problematic. It is often taken for granted and assumed to be the ordinary, basic, straight-forward, unproblematic, unambiguous, context-free, real, true, primary, central or original meaning of a word or expression. A closer look at the issue, however, reveals that these epithets are as vague and poorly defined as the notion which they aim to explain (see e.g. Cacciari 1993, Glucksberg 2001, Leezenberg 2001 or Gibbs 1994 for more discussion).

Gibbs (1994) states that the notion of literal is too elusive, that there is no theory which provides a comprehensive account of literal meaning. He also questions the very possibility of a stable concept of literal meanings for either words or sentences. Nevertheless, his statements like “a theory of idiom processing must await a better idea of what is meant when it is claimed that idioms, or any linguistic construction, can be processed or analyzed literally” (Gibbs 1994: 287) do not prevent Gibbs himself from widely using the term literal in his account on idiomaticity. Glucksberg (2001), on the other hand, suggests that a difference should be made between the folk and the linguistic theory of literal. His position is that the concept of literal cannot be explicitly defined except in terms of formal linguistic-theory (as maximally decontextualized rule-based abstraction). Within folk theory of language, when judging specific examples the distinctions between the literal and the nonliteral become graded, rather than discrete, suggesting that there is a continuum from the literal to the nonliteral (Ortony 1979).

In the present approach, the term literal meaning will be abandoned (except when referring to theories that use this term) as incompatible with the view that conceptual structure is a single level of semantic representation. I assume that there cannot be any principally different meanings – both the literal and the idiomatic structures are formed using the same primitive units and following the same conceptual formation principles. In this respect, there is nothing idiomatic about an idiom’s actual meaning per se; the idiosyncrasies can only be found in the

linking of conceptual structure to another representational module (syntax). But, in order to be able to postulate idiosyncratic linking, one has to accept that there are linguistic phenomena where linking between these two modules is governed by some regular principles. Conceptual Semantics is a theory which both emphasizes the importance of regular structures and does not downplay irregular ones (Nikanne 2005a, 2008a, 2008b). Thus, instead of the term literal meaning I will use the notion of conceptual structure (CS) as a structure which is mapped to the syntactic module according to regular principles of syntactico-semantic linking, as opposed to a structure whose idiosyncratic linking to syntax is licensed by a PU. The latter will be referred to as the CS/PU.

### **3.4.3.2 Parallel conceptual structures in ambiguous PUs**

Idioms that can be assigned both a literal interpretation (in my terms CS) and a figurative one (CS/PU) are called ambiguous (e.g. Colombo 1993, Dobrovolskij & Piirainen 2005) or literally plausible (Titone & Connine 1994). Which conceptual structure has to be included into the formal linguistic description of a PU? To answer this question one has to take into consideration the nature of linguistic processing. There are in fact several different models, each providing a different explanation of how literal vs. idiomatic meaning is processed in idioms. Gibbs (1993: 75) points out that there is no single answer to the question of how idioms are understood: people may not always analyze the literal meaning of idioms during comprehension, but when they attempt to comprehend “certain kinds of idioms” they might also process the individual word meanings. According to the model proposed by e.g. Bobrow & Bell (1973), Weinreich (1969), Clark & Lucy (1975), Janus & Bever (1985), Lyons (1977), Grice (1975) and Searle (1979), the literal meaning of an utterance is always derived first, tested against the context of the utterance, and after that, depending on whether it makes or does not make sense, it is either accepted as the utterance meaning or rejected, and in the latter case an alternative nonliteral meaning is retrieved. Gibbs (1980, 1985, 1986, 1994), on the contrary, claims that idioms are understood directly without any analysis of their literal meanings.

Configuration model, represented by Cacciari & Tabossi (1988), Cacciari (1993) and Tabossi & Zardoni (1993), shows that the literal meaning of idiomatic constituents is activated immediately and the

idiomatic meaning is activated when input is sufficient for the configuration to be recognized as an idiom (the notion of the idiom key refers to the point in an idiomatic string at which the recognition takes place). On the other hand, Estill & Kemper (1982), Swinney & Cutler (1979), Ortony et al. (1978), Stock, Slack & Ortony (1993), McElree & Griffith (1995), McElree & Nordlie (1999), Cutting & Bock (1997) and Colombo (1993) present convincing evidence for the parallel generation and simultaneous accessibility of both literal and idiomatic (figurative) senses of the phrase, even in contexts that bias one of the meanings. Their results speak in favour of the exhaustive access model of activation of meanings in ambiguous lexical units, according to which all meanings are activated automatically and selection of the appropriate meaning is made later. The notion of promiscuous integration of competing structures and their resolution at a later stage also appears in Jackendoff (2002: 211). Swinney (1979) and Tanenhaus, Leiman & Seidenberg (1979) also argue that all of a word's possible meanings are activated when it is first heard in a sentence and are pared down after the word is integrated with the context.

In view of the latter facts, CS has to be included into the formal description of ambiguous PUs alongside the CS/PU. Yet another reason why one would need to do that has to do with syntactic processing. Nikanne (1997a, 2005a) argues that the determination of subject and object is based on the lexical conceptual structure of the verb, and that thematic arguments are not directly linked to syntax. In syntactically well-formed expressions the CS has to be constructed at the stage of integration since it specifies the lexical argument linking, i.e. linking of thematic arguments to syntactic arguments (subject and object) via the intermediate DA level. In so-called non-analyzable PUs thematic arguments that appear in their rule-based argument structure and are linked to syntactic arguments are missing from the argument structure of the CS/PU. For instance, the Finnish PU *X HEITTÄÄ LUSIKKA NURKKAAN* |  $NP_{SUBJ}[N\{X\}] V\{heittää\text{ 'throw'}\} NP_{OBJ}[N_{SG}\{lusikka\text{ 'spoon'}\}] PP_{ILL}[NP_{SG}\{nurkka\text{ 'corner'}\}]$  | lit. 'X casts a spoon into a corner', id. 'X dies' has a variant intransitive construction  $\langle X:N \rangle LUSIKKA LENTÄÄ NURKKAAN$  |  $\langle NP_{GEN}/PP_{ABL}[N\{X\}] \rangle NP_{SUBJ}[N_{SG}\{lusikka\text{ 'spoon'}\}] V\{lentää\text{ 'fly'}\} PP_{ILL}[NP_{SG}\{nurkka\text{ 'corner'}\}]$  | lit. '<X's> spoon flies into a corner', id. 'X dies' (Figure 59 below). The rule-based CS of this construction follows the argument structure of the verb *lentää* 'fly', where the thematic argument

SPOON corresponds via DA1 to the subject. On the other hand, the argument structure of the CS/PU, which is a conceptualization of the event of DYING, does not include SPOON as its thematic argument. Thus, based only on the argument structure of the CS/PU, it would not be possible to determine syntactic arguments in this expression.

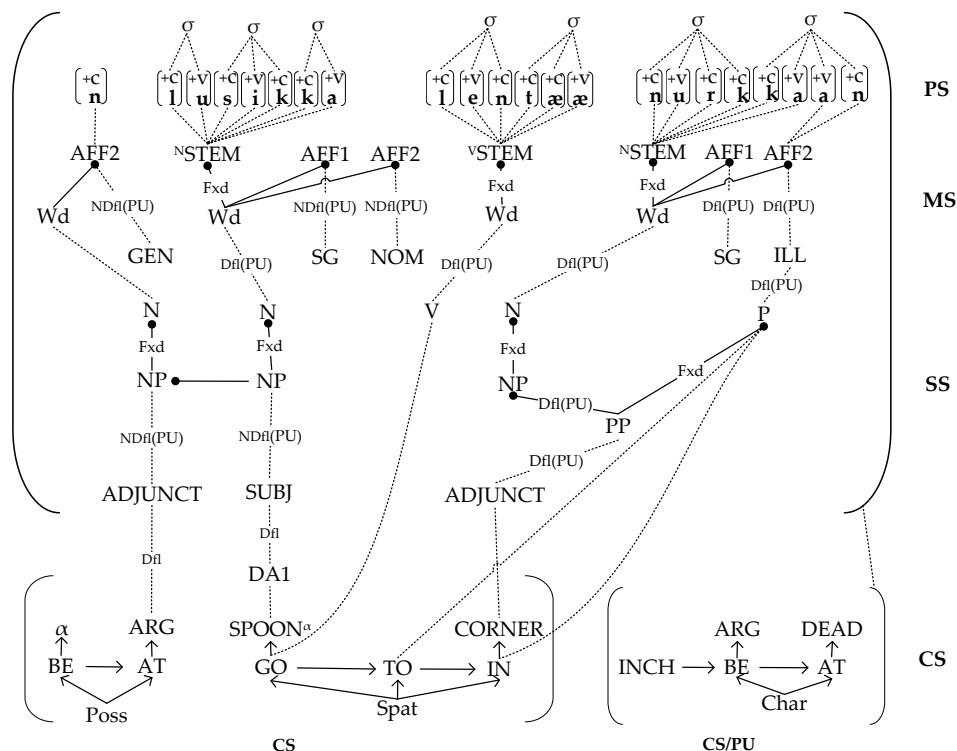


Figure 59 Thematic argument linking to syntactic arguments in a non-analyzable Finnish PU variant <X:N> LUSIKKA LENTÄÄ NURKKAAN lit. '<X's> spoon flies into a corner', id. 'X dies'.

As can be seen in Figure 59, a single correspondence is assumed to exist between PS-MS-SS and the CS/PU as bigger chunks. At the same time, no one-to-one or one-to-many direct correspondence links are drawn between the thematic arguments of the CS/PU on the one hand and syntactic constituents on the other. Moreover, I do not assume that such links can be postulated even in so-called analyzable PUs for a simple reason: since linking to syntactic arguments via DAs is already taken care of by thematic arguments of the rule-based CS, no additional linking to the level of syntax is possible. The problem of analyzability will be tackled in the next section.

### 3.4.3.3 The problem of analyzability

The generally accepted way to treat the analyzability of PUs is to ascribe autonomous idiomatic meanings to their lexical constituents. Analyzability is often described as some kind of isomorphism:

[...] one-to-one correspondence between the formal structure of the expression and the structure of its semantic interpretation, in the sense that there exists systematic correlation between the parts of the semantic value of the expression as a whole and the constituent parts of that expression (Geeraerts 1995<sup>112</sup>: 61).

In other words, this definition implies that there can indeed exist a one-to-one correspondences between thematic arguments and functions of the CS/PU (parts of the structure of semantic interpretation), syntactic constituents and corresponding units of morphophonological structure (constituent parts of the formal structure). However, there are several problems here.

First of all, in Conceptual Semantics even regular rule-based linking is not assumed to be a trivial one-to-one linking; one element of representation may map to one, none or several elements in another representation (Nikanne 2008a). For instance, in Figure 59 above, the morphosyntactic category ILL (illative case) is mapped to two different thematic functions – TO and IN; in Figure 62 below, the Finnish verb *maalata* ‘paint’ is mapped onto the thematic functions CAUSE, GO, TO, ON of the f-chain and the implicit thematic argument [REPRESENTATION OF] on the argument level. Jackendoff (2002: 13-15, 427) also points out that one should not expect isomorphism neither in the syntax-semantics interface nor in the interfaces between syntax and phonology, among phonological tiers, between phonology, acoustics and gesture. Therefore, one cannot in principle expect any one-to-one linking between the CS/PU and other levels of representation either.

Secondly, one cannot expect direct correspondences between the CS/PU conceptual arguments and the syntactic-phonological levels of representation without any principles of correspondence: in the same vein as the linking of CS thematic arguments to syntactic arguments is governed by the intermediate DA-system, CS/PU arguments cannot be just randomly linked to syntax. Since thematic arguments licensed by the

---

<sup>112</sup> For a more detailed review of Geeraerts’ (1995) model of compositionality see Section 2.3.2.



CS/PU are neither potential DAs nor potential syntactic arguments (simply because these roles are already assigned by the CS), there seems to be no way to link them directly to syntax even in so-called semantically analyzable (decomposable) PUs.

What about motivation? Could it provide such linking? Dobrovol'skij & Piirainen (2005) also explore the idea that conventional figurative units possess two conceptual structures – the actual meaning, i.e. the figurative or idiomatic meaning, and the image component, which they define as “the traces of the literal meaning inherited by the figurative meaning” (ibid. 14). In their opinion, the image component mediates between the lexical structure and the figurative meaning. By mediating they primarily mean motivating links. However, since motivation links only indicate correspondences between the two conceptual structures, they cannot establish any direct connections between the thematic arguments of the CS/PU on the one hand, and syntax on the other. In this respect, the CS indeed mediates between the syntactic structure and the CS/PU (in motivated PUs), as Dobrovol'skij & Piirainen (2005) suggest. Nevertheless, they remark that motivation cannot be equated with and does not necessarily affect a unit's analyzability, rather the analyzability is one of the reasons why a PU is perceived as being motivated (ibid. 84).

Apart from the assignment of independent subidiomatic meanings to PU constituents, there have been attempts to describe analyzability in terms of reference. The idiomatic referentiality of components is mentioned by e.g. Gibbs & Nayak (1989) and Gibbs (1994), who establish the difference between normally decomposable and abnormally decomposable idioms based on the relationship of individual components to their idiomatic or figurative referents: in normally decomposable idioms “components have a more direct relation to their figurative referents” than in abnormally decomposable idioms, where “each part does not *by itself* refer to some component of the idiomatic referent but only to some metaphorical relation between the individual part and the referent”, e.g. *question* in *POP THE QUESTION* is assumed to directly refer to a marriage proposal, as well as *tongue* and *lip* in *HOLD YOUR TONGUE* and *BUTTON YOUR LIP* directly refer to speech; while *torch* in *CARRY THE TORCH FOR SOMEBODY* metaphorically refers to warm feelings (Gibbs 1994: 279, 281). Nenonen (2002: 20) also remarks that the Finnish PUs *OTTAA HUIKKAA* ‘drink alcohol’ and *OTTAA LÄRVIT* ‘drink a lot of alcohol’ are isomorphic, because both of their component words have a referent,

although the relationship is not directly literal. Unfortunately, the notion of direct and less direct, metaphorical reference is presented in Gibbs (1994) and Gibbs & Nayak (1989) only vaguely, e.g. it is not clear what the authors exactly mean by the component of the referent, or the fact that an individual part can refer to a metaphorical relation between itself and its referent. One can also argue that the relationship between the *question* and the marriage proposal, or *tongue* and *lip* and speech is not direct, but is rather based on metonymy.

The latter account on analyzability as PU referentiality of constituents is worthwhile exploring in more detail. If I assume that linking of the CS/PU to the level of syntactic representation is mediated by the assignment of PU referents to the syntactic constituents of the PU, I would need to include into my formal description of PUs a separate level of representation which encodes referential claims for both the CS and the CS/PU. In Section 3.4.3.4 Jackendoff's conceptualist theory of reference, which has been earlier described in Section 3.2.9, will be adapted for the purposes of the Tiernet model-based PU analysis.

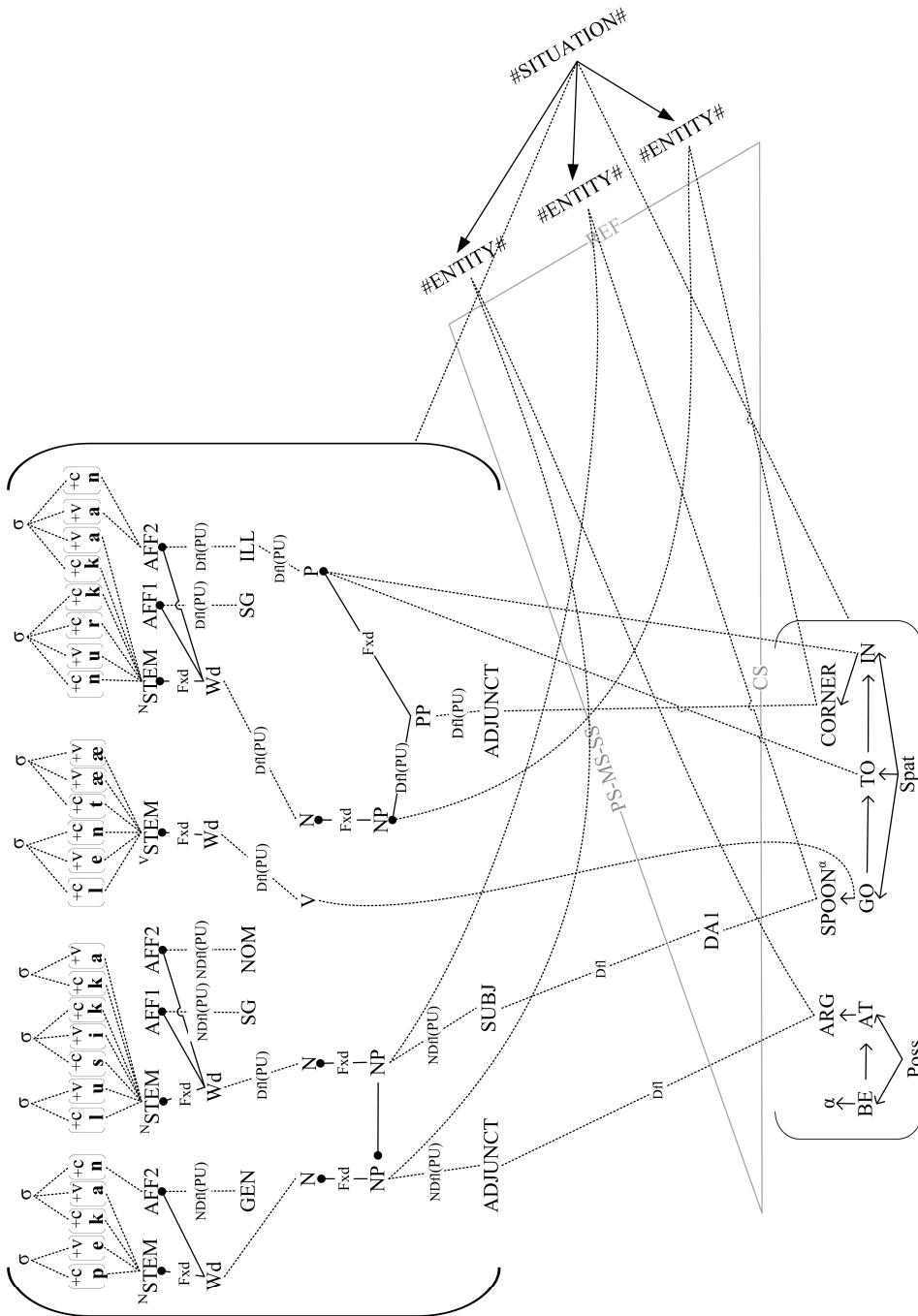
#### 3.4.3.4 Reference in PUs

A crucial part of semantic description is a formalization of the referential claims which a linguistic expression makes about certain entities. Since my goal is to formalize the PU's internal structure, the latter has to include a separate level of representation which explicates referential relations that underlie establishing the contextualized meaning of the expression. One will need an interface, a point of interconnection between our conceptualization of the PU and our conceptualization of the context, or, to be more specific, of those elements of the context which can be established as co-referential with the elements of the PU's structure. Some previous theoretical studies concerning the semantic relations that might exist between the PU and its context have been presented in Section 2.4 of Chapter 2 and criticized for their lack of explicitness. In the present section I will discuss a possible application of Conceptual Semantics to the problem of reference in PUs.

Figure 60 below is an attempt to incorporate the referential tier into the Tiernet model. It presents an analysis of a referring regular linguistic expression *Pekan lusikka lentää nurkkaan* 'Pekka's spoon flies into the corner' (the formal description of the Finnish intransitive PU construction  $X:N \text{ LUSIKKA LENTÄÄ NURKKAAN} \mid NP_{GEN}\{X\} > NP_{SUB}\{\textit{lusikka 'spoon'}\} V\{\textit{lentää}$

*'fly'}*  $PP_{SG\ ILL}[NP\{nurkka\ 'corner'}]\mid$  lit. '<X's> spoon flies into a corner; id. X dies' is yet to come in Figure 61).

Here, references to the conceptualized world entities are surrounded by ## (this metalanguage notation is taken from Jackendoff 1983, where he uses it to denote projected-world entities). The satisfaction of the expression is not taken into consideration here, thus the purported referents may or may not be present in the hearer's f-knowledge base or readily available context. The indexical features of the referential tier are marked as #SITUATION# or #ENTITY#. Correspondence links between indexical features and syntactic constituents of the linguistic expression encode referential commitments to these constituents. In Jackendoff's (1972; 2002: 326) theory, the intended reference of a declarative sentence S is a conceptualized #SITUATION#. However, since the syntactic category S is abandoned from my description as not being truly compatible with the network model, the correspondence link is drawn from the #SITUATION# to a larger fragment of the network enclosed in double brackets – the PS-MS-SS area of the triangle, i.e. the phonological structure (PS), the morphological structure (MS) and the syntactic structure (SS). The absence of any direct correspondence between individual nodes of phonology and referential indices is in accord with Jackendoff's view of the syntax-semantics-context linking, which does not include "interface rules that directly relate contextual understanding to phonological structure", at the same time assuming that there are "interfaces linking them through semantics and syntax" (Jackendoff 2007a: 68).



To summarize, correspondence between indexical features and syntactic constituents of the expression in Figure 60 is as follows:

- Indexical features of the three conceptualized #ENTITIES# correspond to the NPs – *NP{Pekka}*, *NP{lusikka 'spoon'}* and *NP{nurkka 'corner'}*.
- The index of the conceptualized #SITUATION# corresponds to the PS-MS-SS area as a whole.
- Indexical features are also mapped by correspondence links to the CS of the expression:
  - the #ENTITY# referents correspond to the thematic arguments [PEKKA], [SPOON] and [CORNER]; and
  - the #SITUATION# referent corresponds to the CS as a whole.

Apart from incorporating the referential tier into my network description, I have several other important methodological tasks. Since my goal is to demonstrate the difference which exists in the linking between the reference tier and other tiers in PUs as opposed to regular expressions, I first need to formulate general rule-based principles of such linking and then compare them to those, which are licensed by PUs. Jackendoff assumes that in natural language referentiality is the unmarked case, while nonreferential use of a phrase in a sentence can be traced to an explicit lexical or grammatical marker. I start from the assumption that the distribution of referential phrases in natural language follows the Referentiality Principle (Jackendoff 1983: 69-70):

Unless there is a linguistic marking to the contrary, all phrases that express conceptual constituents are referential.

Some principles of mapping between the CS and the referential tier of a declarative sentence are formulated in (A) – (E) below:

- (A) The conceptual structure of a SITUATION (EVENT or STATE) corresponds to the referential index of a #SITUATION#.
- (B) Theta-arguments required by functions of the f-chain are potential referents and are mapped to the referential tier indices of the conceptualized #ENTITIES# (these can belong to different ontological types; see Jackendoff 2002b: 315-318) – participants of the situation.
- (C) Implicit theta-arguments are not potential referents and are not mapped onto the referential tier.
- (D) If a theta-argument is bound by another theta-argument, they are mapped onto the same index on the referential tier.
- (E) If the f-chain f1 and f2 functions are selected by the Characterizing (Identificational) semantic field, thematic argument (TYPE or

PROPERTY) selected by f1 is not a potential referent and is not mapped onto the referential tier<sup>113</sup>.

While these general principles are valid for regular expressions, PUs do not exactly follow all of them. Let us compare the formal description of the sentence in Figure 60 above to the description of the Finnish intransitive construction  $X:N$  *LUSIKKA LENTÄÄ NURKKAAN* |  $\langle NP_{GEN}\{X\} \rangle NP_{SUBJ}\{lusikka \text{ 'spoon'}\} V\{lentää \text{ 'fly'}\} PP_{SG\ ILL}[NP\{nurkka \text{ 'corner'}\}]$  | lit. '<X's> spoon flies into a corner; id. X dies' – a variant of the Finnish PU  $X$  *HEITTÄÄ LUSIKKA NURKKAAN* |  $NP_{SUBJ}[N\{X\}] V\{heittää \text{ 'cast'}\} NP_{OBJ}\{lusikka \text{ 'spoon'}\} PP_{ILL}[NP_{SG}\{nurkka \text{ 'corner'}\}]$  | lit. 'X casts a spoon into a corner; id. 'X dies' in Figure 61 below.

The logical consequence of including two parallel conceptual structures (CS and CS/PU) into the formal description of PUs will be to assume their referential ambiguity. Since the two conceptual structures corresponding to the same SS encode information about completely different SITUATIONS ('the transition of someone's spoon into the corner' and 'someone's becoming dead'), each conceptual structure as a whole corresponds on the referential tier to the indexical of different conceptualized #SITUATION#, while the SS corresponds to both indices. The notation #SITUATION#/PU indicates that this particular indexical is licensed by the PU, as opposed to the conceptualized #SITUATION# of the rule-based CS. The Referentiality Principle (Jackendoff 1983: 69-70), according to which all phrases that express conceptual constituents are referential, is valid for the CS in the given example:  $NP_{GEN}\{Pekka\}$ ,  $NP\{lusikka \text{ 'spoon'}\}$  and  $NP\{nurkka \text{ 'corner'}\}$  indeed correspond to thematic arguments [PEKKA], [SPOON] and [CORNER], on the one hand, and to referential #ENTITY# indices, on the other. However, the same principle cannot be applied to the CS/PU, since, as I have earlier argued, its conceptual constituents have no direct correspondence to syntax. Thus, the #SITUATION#/PU referentiality of syntactic constituents cannot be inferred on the basis of their linking to the CS/PU, all the more so if one assumes that this linking is mediated by the referential tier.

---

<sup>113</sup> On a syntactic level this argument corresponds to a predicative in a sentence. Adjectives and predicate NPs do not carry referential claims, they contribute only to descriptive features but no new individual (Jackendoff 2002: 396).



How can the PU referentiality be established? The most clear cases are represented by the so-called regular slots (Moon 1998: 98), i.e. lexically unspecified syntactic constituents, which are licensed by the construction and are to be filled according to context (e.g. the one indicated by *X* in *X:N LUSIKKA LENTÄÄ NURKKAAN* |  $NP_{GEN}\{X\} > NP_{SUB}\{lusikka \text{ 'spoon' }\} V\{lentää \text{ 'fly' }\} PP_{SG\ ILL}[NP\{nurkka \text{ 'corner' }\}]$  | lit. '<X's> spoon flies into a corner; id. X dies'). A regular slot syntactic constituent is mapped to an unspecified thematic argument in the CS according to regular correspondence principles described in Section 3.2.8 (i.e. the subject argument by default is mapped to DA1 and the object to DA2), while the position of a co-referential thematic argument in the CS/PU is licensed by the latter structure. Both thematic arguments and the syntactic constituent itself correspond on the referential tier to the same index of a conceptualized #ENTITY#, which appears as a character in both the unmarked #SITUATION# and the #SITUATION#/PU, i.e. the #ENTITY#, which possesses the 'spoon' in the unmarked #SITUATION# is the same #ENTITY# which becomes DEAD in the #SITUATION#/PU.

What about other conceptual constituents of the CS/PU? According to the principle formulated in (E) earlier in this section, [DEAD] is not a potential referent since it is a PROPERTY thematic argument selected by *f1*, while the *f1* and *f2* functions are selected by the Characterizing semantic field. Thus, since the only available character index in the #SITUATION#/PU is taken by *X*, the remaining NP constituents *lusikka* 'spoon' and *nurkka* 'corner' are left without any possible linking node on the referential tier. Therefore, these constituents can be regarded as #SITUATION#/PU non-referential.

Constituents of the Finnish PU  $NP_{SUB}[N\{X\}] MAALAA PIRUJA SEINILLE$  <Y:STÄ> 'lit. X paints devils onto walls <about Y>, id. X exaggerates potential risks <caused by Y>' (Figure 62 below) demonstrate a different kind of referentiality<sup>114</sup>.

---

<sup>114</sup> In order to avoid an excessive tangle of links in the picture, the correspondences between syntactic constituents, conceptual constituents of the LCS and REFs of the unmarked #Situation# are indicated by lower case presubscript letter indices.



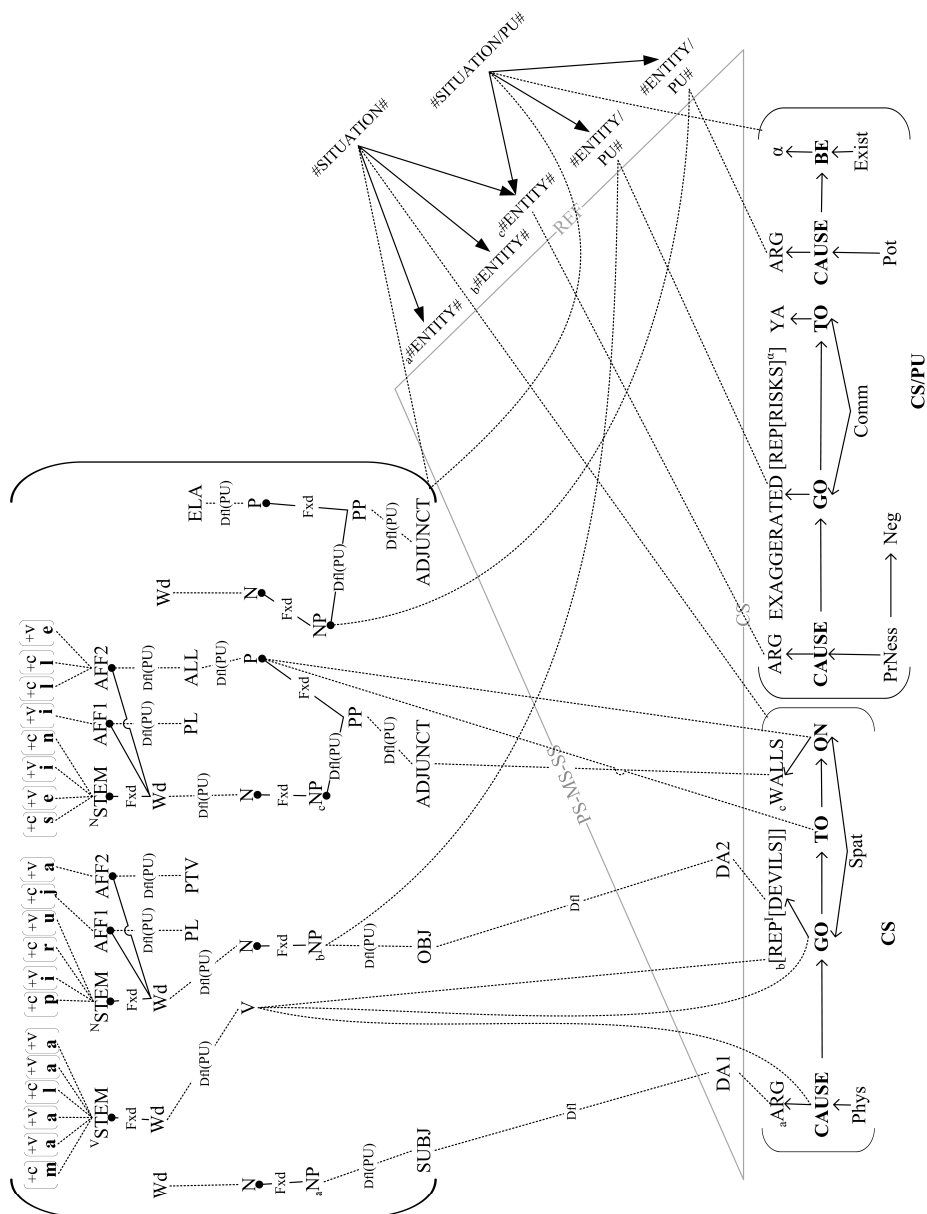


Figure 62 The Finnish PU NP<sub>SUBJ</sub>[N{X}] MAALAA PIRUJA SEINILLE <Y:STÄ> 'lit. X paints devils onto walls <about Y>, id. X exaggerates potential risks <caused by Y>'

Apart from the regular slot grammatical subject of the sentence, #SITUATION#/PU referentiality can be ascribed to the object NP *piruja* ‘devils’. As is seen in Examples (74) – (76) below, this constituent can be taken to refer to the potential [RISKS]:

- (74) *Hyvä esimerkki markkinointitavasta, jossa maalataan pahimmat mahdolliset pirut seinille, eikä kuitenkaan viitata minkäänlaiseen tutkimukselliseen tietoon.*<sup>115</sup>  
lit. ‘A good example of a marketing practice, where **they paint the worst possible devils on walls**, but still do not refer to any kind of research data.’
- (75) *Emme siis maalaile piruja seinille, vaan pikemminkin näytämme missä ne pirut varsinaisesti ovat.*<sup>116</sup>  
lit. ‘So we do not **paint devils on walls**, but we rather show where **the devils** truly are.’
- (76) *Eivät ne ole seinälle maalattuja piruja, Kassakaappi. Todellisia piruja ne ovat, avaa silmäsi, lue uutisia.*<sup>117</sup>  
lit. ‘These are not **devils painted on walls**, Kassakaappi. These are **real devils**, open your eyes, read the news.’

Both [DEVILS] and [RISKS] are thematic arguments selected by the same zone 2 function GO in CS and the CS/PU. Although the matter requires a more extensive study of formalized structures of a greater number of PUs than can be done within the scope of this study, there is a good reason to believe that analogy plays an important role in the assignment of #SITUATION#/PU referents to syntactic constituents. I.e., given the identity of f-chains in both conceptual structures, there is a high probability that syntactic constituents, corresponding to thematic arguments in the CS, will be mapped onto the same #ENTITY# indices in the #SITUATION#/PU, as arguments selected by the same thematic functions in the CS/PU. Consequently, except for regular slots, arguments which are not selected by the same function, cannot in principle be expected to map onto the same #SITUATION#/PU referent.

In the PU  $NP_{SUBJ}[N\{X\}]$  MAALAA PIRUJA SEINILLE <Y:STÄ> ‘lit. X paints devils onto walls <about Y>, id. X exaggerates potential risks <caused by Y>’, f-chains in the CS (CAUSE→GO→TO→ON) and in the CS/PU (CAUSE→GO→TO) are similar in zone 3 and zone 2 (CAUSE→GO), but differ in zone 1. In the CS the zone 1 function TO does not take any

<sup>115</sup> <http://www.rintamamiestalo.fi/viewtopic.php?p=46143#p46143>, Mar 17 2008

<sup>116</sup> [http://www.vertigo.cd/index.php?option=com\\_content&task=view&id=2673&Itemid=27](http://www.vertigo.cd/index.php?option=com_content&task=view&id=2673&Itemid=27), Aug 5 2008

<sup>117</sup> <http://keskustelu.plaza.fi/viesti/9468118/29/>, Oct 13 2008

argument, but selects another function ON, which in its turn selects the landmark – the argument WALLS; while in the CS/PU, function TO selects a generic argument YA (generic individual, to whom the exaggerated representation of the risks is communicated; see Jackendoff (2008: 214) for the explanation of the term YA). However, the NP constituent *seinät* ‘walls’ does not seem to refer to this generic character. In Examples (77) – (79) below modifiers *meidän* ‘our’, *Nokian ja Siemensin* ‘Nokia’s and Siemens’ and *kaikki* ‘all’ rather refer to the protagonist who is at risk of some bad outcome (this thematic argument is not present in the formal description of the CS/PU in Figure 62, but is assumed to be a part of the LCS of the lexical unit *risk*).

- (77) *Vaikka Merja maalasikin valmiiksi piruja meidän seinille, niin siitä huolimatta meillä sujui Millan kanssa kaikki todella hienosti, mitään hankaluuksia ei ollut.*<sup>118</sup>

lit. ‘Although Merja had **painted devils on our walls**, nevertheless everything went really fine for us with Milla, there were no difficulties.’

- (78) *Forrester maalaa piruja Nokian ja Siemensin seinille*

*Arvostetun tutkimusyhtiö Forrester Researchin teleanalyttikko Lars Godell kommentoi Nokian ja Siemensin päätöstä yhdistää verkkotoimintansa. Godell on epäileväinen kaupan hyötyjen suhteen ja esittää useita teorioita siitä, mikä voi mennä pieleen.*<sup>119</sup>

lit. ‘Forrester **paints devils on Nokia’s and Siemens’ walls**

Telco analyst of the respected research company Forrester Research Lars Godell commented on Nokia’s and Siemens’ decision to unite their network activity. Godell is doubtful about the benefits of the deal and presents several theories on what can go wrong.’

- (79) *Kaikilla seinillä ei näy piruja.*<sup>120</sup>

lit. ‘**Devils** are not seen **on all walls**.’

Finally, the optional regular slot NP<sub>ELA</sub> constituent, which is co-referential with the argument of CAUSE (the potential causer of risks) in the second conceptual clause of the CS/PU, cannot be mapped onto the CS, since its dominating node PP<sub>ELA</sub> does not correspond to any argument selected by the lexical f-chain of the verb *maalata* ‘paint’. In Nikanne’s (2005a) notation this relation can be formalized as V -/> PP<sub>ELA</sub>, which

<sup>118</sup> <http://eveliah.suntuubi.com/?cat=8&y=2008&m=2>, Aug 3 2008

<sup>119</sup> <http://www.digitoday.fi/mobiili/2006/06/20/forrester-maalaa-piruja-nokian-ja-siemensin-seinille/20068548/66>, Jun 20 2006

<sup>120</sup> <http://vanha.verkkouutiset.fi/arkisto/paakirjoitus/138762.html>, Nov 27 2008

stands for ‘PP<sub>ELA</sub> is not selected by V directly or indirectly’. Thus, on the referential tier the constituent cannot be mapped to any index of the unmarked #SITUATION#, but refers only to an #ENTITY#/PU in the #SITUATION#/PU.

### 3.4.3.5 Unspecified thematic arguments and context dependency in the Finnish PU *HEITTÄÄ HELMIÄ SIOILLE* ‘throw pearls to pigs’

Figure 63 below shows the Finnish PU *X HEITTÄÄ HELMIÄ SIOILLE* | *NP<sub>SUBJ</sub>[N{X}] V{heittää ‘throw’} NP<sub>OBJ</sub>[N<sub>PL</sub>{helmi ‘pearl’}] PP<sub>ALL</sub>[NP<sub>PL</sub>{sika ‘pig’}]* | lit. ‘X throws pearls to pigs’, id. ‘X causes a transfer of some entity, evaluated by the speaker as good, to a recipient, evaluated as inadequate in some way’, which has identical f-chains in both the CS and the CS/PU. All NP constituents of this PU are #SITUATION#/PU referential, i.e. they are all mapped onto #ENTITY#/PU referential indices in the scope of the #SITUATION#/PU. Both in the CS and in the CS/PU these indices correspond to thematic arguments selected by the same functions. A regular slot *NP<sub>SUBJ</sub>[N{X}]* constituent corresponds to the referential index shared by both #SITUATIONS# and its thematic argument counterpart is selected by the same function CAUSE in both conceptual structures. The NPs *helmi* ‘pearl’ and *sika* ‘pig’ have respective #ENTITY# referential indices in both #SITUATIONS#. Their #ENTITY#/PU REFs in turn correspond in the CS/PU to the two unspecified thematic arguments, which are selected by the same thematic tier functions (GO and TO) as the conceptual arguments [PEARL] and [PIG] in the CS (THEME and LANDMARK positions).

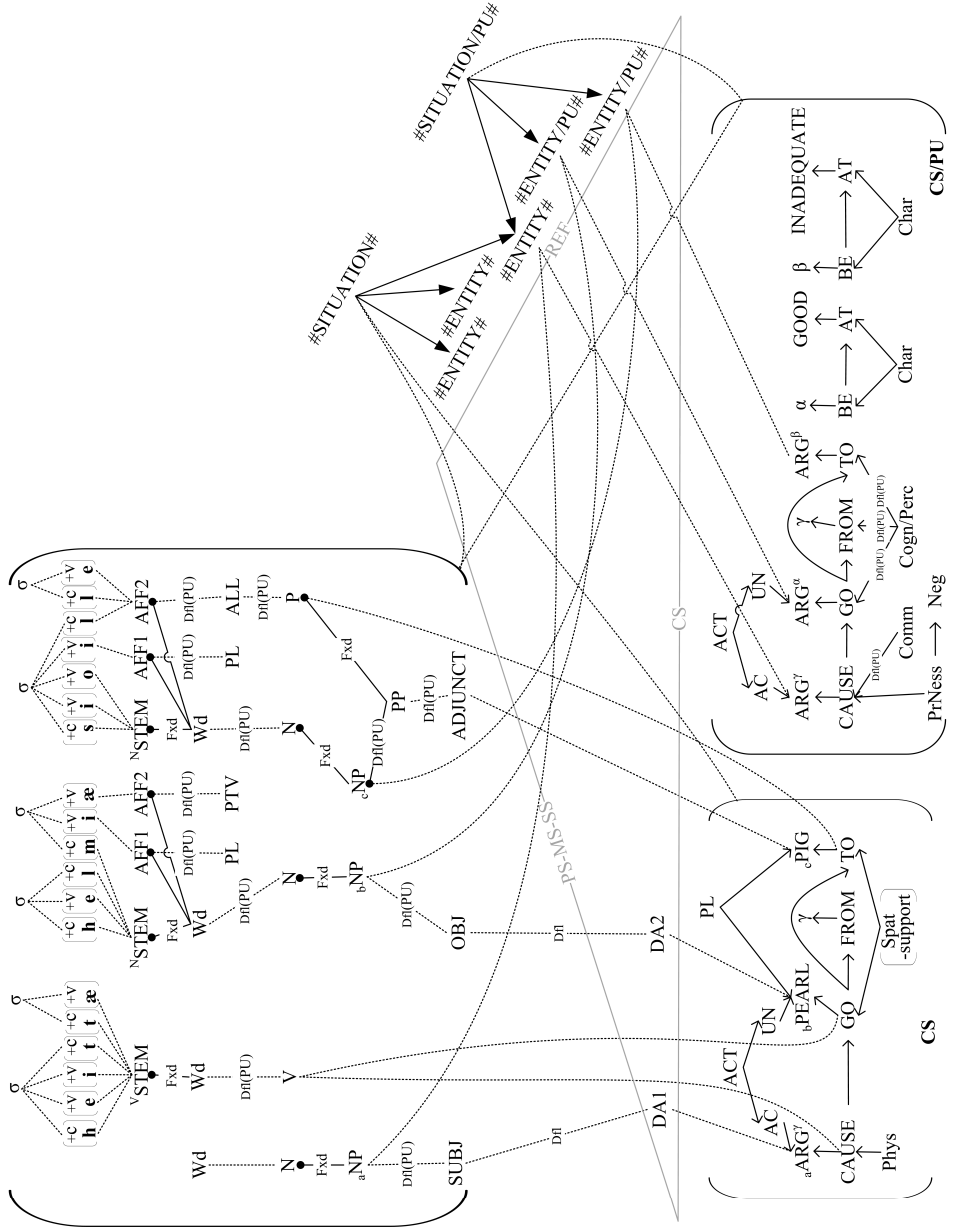


Figure 63 The Finnish PU X HEITTÄÄ HELMIÄ SIOILLE |  $\text{NP}_{\text{SUBJ}}[\text{N}\{\text{X}\}]$   $\text{V}\{\text{heittää 'throw'}\}$   $\text{NP}_{\text{OBJ}}[\text{N}_{\text{PL}}\{\text{helmi 'pearl'}}\}]$   $\text{PP}_{\text{ALL}}[\text{NP}_{\text{PL}}\{\text{sika 'pig'}}\}]$  | lit. 'X throws pearls to pigs'

As can be seen from Figure 63 above, none of the thematic arguments of the CS/PU is specified. This makes this PU a semantically ambiguous and context-sensitive unit, which largely underdetermines the proposition expressed. Its disambiguation and specification demand for its referential indeterminacies to be resolved. In order to arrive at the complete truth-conditional content in the form of a specific CS/PU, the respective referential indices have to be derived from the f-knowledge base, including the actual context in which the PU is used. I.e., one has to find the referents of the  $NP_{SUBJ}[N\{X\}]$ , *pearls* and *pigs* in the context. It is the task of REFs to point out the entities of the conceptualized world, which are intended to appear in the unspecified argument slots on the thematic argument tier of the CS/PU.

The CS/PU contains two characterization structures, which predicate certain evaluative PROPERTIES to the arguments  $[\alpha]$  and  $[\beta]$  bound by unspecified arguments  $[ARG]^\alpha$  and  $[ARG]^\beta$  of the main conceptual clause. Since  $[\alpha]$  and  $[\beta]$  are co-referential with  $[ARG]^\alpha$  and  $[ARG]^\beta$ , they are also co-referential with the respective NP constituents *helmet* ‘pearls’ and *siat* ‘pigs’. In the first embedded conceptual clause  $[\alpha]$  is characterized as [GOOD] and in the second one  $[\beta]$  is characterized as [INADEQUATE]<sup>121</sup>. Thus, in order to be indexed as a referent of  $[ARG]^\alpha$  and  $[ARG]^\beta$  and introduced into the argument level of the CS/PU, a context-derived argument candidate has to match its respective property. Consequently, for a successful reference assignment, the evaluative elements themselves are to be highly expected to appear in the context and specify the otherwise abstract properties. This is especially relevant for the PROPERTY of  $[\beta]$ , which, depending on what kind of an adequate behaviour is expected from the referent of  $[ARG]^\beta$  as a reaction to  $[ARG]^\alpha$ , can mean a different kind of inadequacy. These options are summarized in Figure 64 below, the default one is given in boldface. NOT is an operator which selects as its scope a complex predicate, which in turn selects an argument ( $\alpha$ ).

---

<sup>121</sup> Motivational correspondences of these properties to the Evaluation tier are explained in Section 3.4.1.3.3 of this chapter.

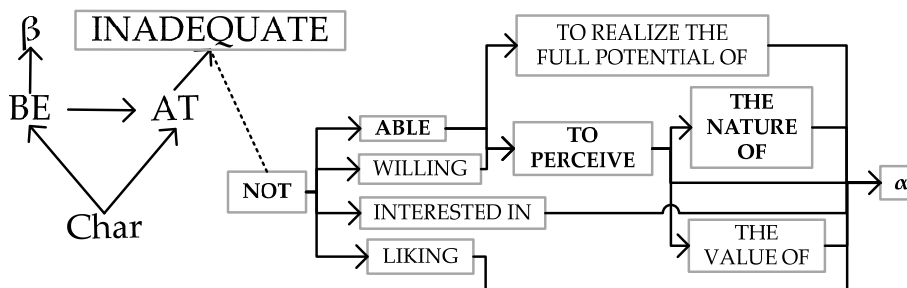


Figure 64 Different realizations of the property INADEQUATE in the CS/PU of HELMIÄ SIOILLE constructions<sup>122</sup>

I will return to a more detailed usage-based analysis of default and non-default linking within the CS/PU and between the CS/PU and the #SITUATION#/PU in Chapter 5. For the present, the following examples (80), (81) and (82) will serve as a brief illustration of the way in which the formal description of reference in the PU *X HEITTÄÄ HELMIÄ SIOILLE* lit. ‘X throws pearls to pigs’ can be applied to actual usage events. In Figure 65, Figure 66 and Figure 67 parts of the context are co-indexed with respective REFs, thus explicating cohesive linking between the text and the idiom’s structure.

Although the number of potential referents in the #SITUATION#/PU is three, it is not necessarily always the case that all of them will be present in context. For instance, in Examples (80) and (82) there is no direct reference to the regular slot  $NP_{SUBJ}[N\{X\}]$  constituent and its argument is represented by a generic YA.

- (80) *Minusta [vika on sinun<sub>2</sub> suodatuksessasi, joka muuntaa kaiken totuudenkin<sub>1</sub> virheelliseksi ja tiedon<sub>1</sub> valheeksi]<sub>5</sub> ... Siten [sinulle<sub>2</sub> oikean<sub>4</sub> tiedon<sub>1</sub> kertominen]<sub>3</sub> on sama kuin heittäisi helmiä sioille.*<sup>123</sup>

lit. ‘I think that [the fault is in your<sub>2</sub> filtering, which changes every truth<sub>1</sub> into false and knowledge<sub>1</sub> into a lie]<sub>5</sub> ... Thus [telling true<sub>4</sub> knowledge<sub>1</sub> to you<sub>2</sub>]<sub>3</sub> is the same as if **throwing pearls to pigs**.’

<sup>122</sup> TO PERCEIVE THE VALUE OF paraphrases the meaning of *appreciate*, and TO PERCEIVE THE NATURE OF paraphrases the meaning of *understand*.

<sup>123</sup> GG: sfnet.keskustelu.uskonto.kristinusko/msg/483ef9a5cb36631e, May 24 2004





- (81) *Sinä<sub>3</sub> olet kirjoittanut, että Ø<sub>3</sub> [et ole kiinnostunut minun<sub>1</sub> ajatuksistani<sub>2</sub>]<sub>5</sub>. [Ei]<sub>4</sub> minun<sub>1</sub> [kannata]<sub>4</sub> heittää helmiä sioille.<sup>124</sup>*  
 lit. ‘You<sub>3</sub> have written that (you<sub>3</sub>) are [not interested in my<sub>1</sub> thoughts<sub>2</sub>]<sub>5</sub>. I<sub>1</sub> [should not]<sub>4</sub> **throw pearls to pigs.**’

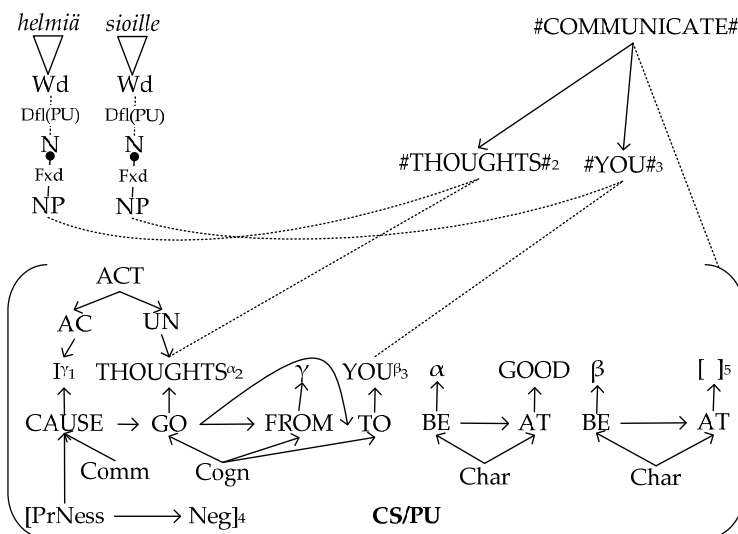


Figure 66 Analysis of the CS/PU in Example (81)

In the following Example (82) (Figure 67) *pearls* are co-referential with *elokuvan kuvasommittelu ja äänisuunnittelu* ‘graphic and sound design of a film’ (GRAPHIC & SOUND ...<sub>1</sub>; #GRAPHIC & SOUND ...#<sub>1</sub>) and *pigs* with a generic person (YA<sub>2</sub>), who goes to the cinema theatre to watch a film, but is unable to appreciate its sound and picture quality because, figuratively speaking, *[hänellä<sub>2</sub> on silmät puusta ja korvat tuohesta]<sub>4</sub>* ‘he has eyes made of wood and ears made of birch bark’<sub>4</sub>. The context also includes a contextual reference to the positive characterization of REF<sub>1</sub> as being [a result of long work]<sub>3</sub>. The regular subject slot is here, too, filled by the generic YA.

- (82) *Täytyy ihmetellä, miksi ihminen<sub>2</sub> ylimalkaan menee elokuvateatteriin, jos [hänellä<sub>2</sub> on silmät puusta ja korvat tuohesta]<sub>4</sub>? Eikö hän<sub>2</sub> voisi tyytyä pysyttelemään tynnyrissään? [Elokuvan kuvasommittelu ja äänisuunnittelu]<sub>1</sub> on [pitkällisen työn tulosta]<sub>3</sub>. Siinä totisesti annetaan helmiä sioille oikein sylrikaupalla.<sup>125</sup>*  
 lit. ‘I can only wonder why a person<sub>2</sub> goes to the cinema at all, if [he<sub>2</sub> has eyes made of wood and ears made of birch bark]<sub>4</sub>? Couldn’t he<sub>2</sub> be content

<sup>124</sup> GG: sfnet.keskustelu.filosofia/msg/bdf29b56c9c77e80, Apr 22 2006

<sup>125</sup> GG: sfnet.harrastus.elokuvat/msg/c2b7bcd49042e32a, Mar 27 2007

with staying in his barrel? [Graphic and sound design of the film]<sub>1</sub> is [a result of long work]<sub>3</sub>. There an armful of **pearls is given to pigs**.'

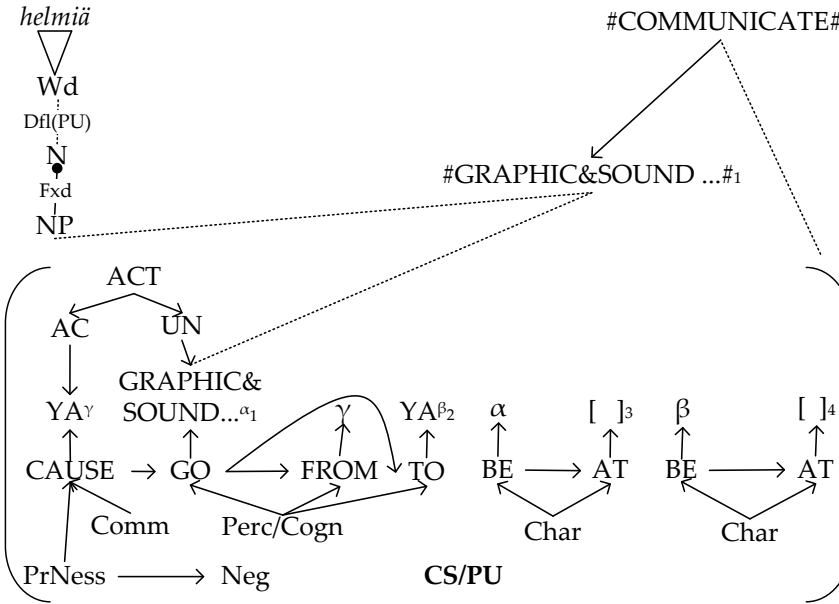


Figure 67 Analysis of the CS/PU in Example (82)

### 3.4.3.6 Some remarks on the semantic field tier (S-tier) and the modal tier in PUs

#### 3.4.3.6.1 The S-tier

The difference between semantic fields in the CS and the CS/PU can be regarded as one of the factors which contribute to the overall idiomaticity of the structure. Since the spatial semantic field is considered as primary in the theory of Conceptual Semantics, I will look at a few examples of a PU where spatial situations conceptualized by the CS correspond to other semantic fields in the CS/PU. For instance, a spatial field in the Finnish PU *X JOUTUU OJASTA ALLIKKOON* lit. 'X ends up from the ditch into the puddle', id. 'X goes from bad to worse', as well as in its English counterpart *X FALLS OUT OF THE FRYING PAN INTO THE FIRE*, corresponds to a circumstantial field in the CS/PU (Figure 68 below). Similar correspondences between the circumstantial field in the CS/PU and the spatial field in the CS can be observed in several other Finnish PUs with the meaning 'X is in a difficult situation, predicament' e.g.: *X ON PINTEESSÄ* 'X is in the clamp', *X ON AHTAALLA/AHDINGOSSA* 'X is in the

narrow', *X ON PUUN JA KUOREN VÄLISSÄ* 'X is between the tree and the bark' etc.<sup>126</sup>

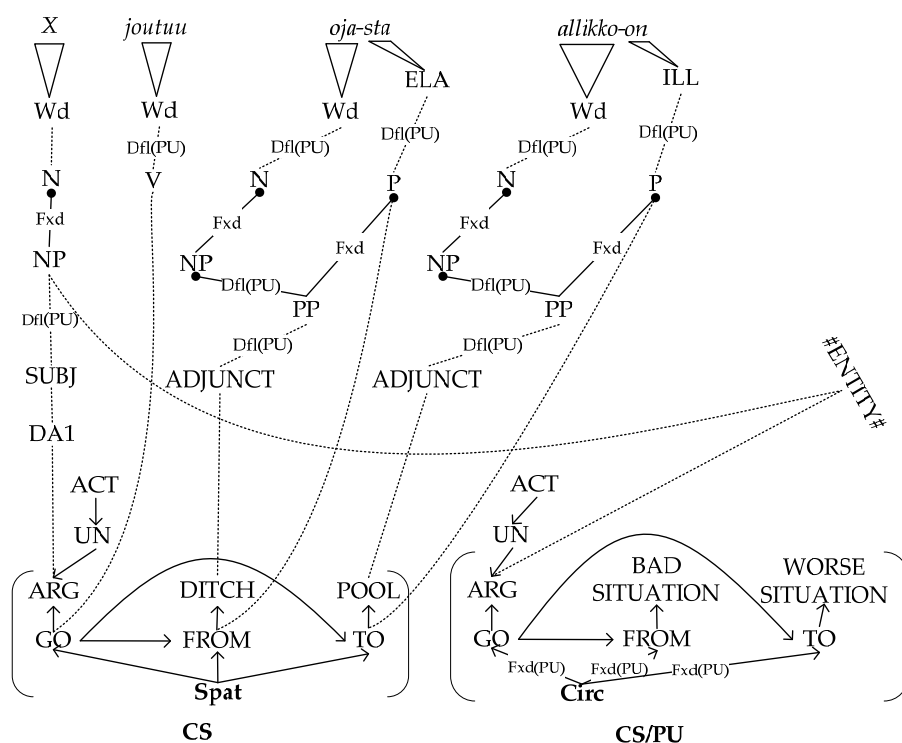


Figure 68 The Finnish PU *X JOUTUU OJASTA ALLIKKOON* lit. 'X ends up from the ditch into the puddle', id. 'X goes from bad to worse'

The spatial field in the CS can also correspond to the characterizing semantic field in the CS/PU, e.g. in the Finnish PUs *X ON SEITSEMÄNNESSÄ TAIVAASSA* lit. 'X is in seventh heaven' (Figure 69 below) and *X ON ONNENSA KUKKULOILLA* lit. 'X is on the hills of his/her happiness' the regular slot subject argument *X* is characterized as EXTREMELY HAPPY in the CS/PU, while in the CS the function BE, which selects this argument, is selected by the spatial field.

<sup>126</sup> The Cognitive Theory of Metaphor would analyze these PUs as instances of the conceptual metaphor DIFFICULTIES ARE IMPEDIMENTS TO MOTION. As Dobrovol'skij & Piirainen (2005: 143) remark, apart from giving a rather abstract metaphoric model, the theory does not provide any tools of analysis, which could e.g. explain semantic differences between the PUs. Other points of criticism against the Cognitive Theory of Metaphor are presented in Chapter 2, Section 2.3.3.

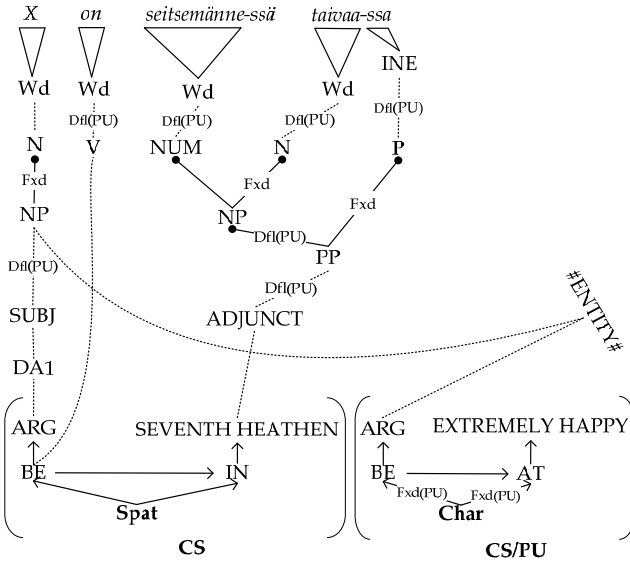


Figure 69 The Finnish PU *X ON SEITSEMÄNNESSÄ TAIVAASSA* lit. 'X is in seventh heaven', id. 'X is extremely happy'

PUs in Figures 68 and 69 specify fixed semantic fields in their CS/PU. But it is not necessarily the case that a PU's conceptual structure always features only one invariant semantic field. The Finnish PU *X HEITTÄÄ HELMIÄ SIOILLE* |  $NP_{SUBJ}[N\{X\}]$   $V\{heittää\}$  'throw'  $NP_{OBJ}[N_{PL}\{helmi\}$  'pearl']  $PP_{ALL}[NP_{PL}\{sika\}$  'pig']] | lit. 'X throws pearls to pigs', id. 'X causes a transfer of some entity, evaluated by the speaker as good, to a recipient, evaluated as inadequate in some way' and its construction family demonstrate considerable variation in the linking to the semantic field. Since the objective of my analysis is to arrive at a semantic representation, which is compatible with all possible contextual realizations, I have to somehow deal with this variation. This could be achieved in two ways: either by leaving the S-tier in the CS/PU completely unspecified and assuming that it can be used in different semantic fields or by recording (on the basis of corpus evidence) the default status of one semantic field, at the same time admitting the possibility of others. I have chosen the latter of these two solutions: in the CS/PU presented in Figure 63 the S-tier of zone 3 features the default Communicative semantic field, while the S-tier of zones 2 and 1 features the default Cognitive/Perceptual semantic fields. Just like the unspecified thematic arguments, the semantic fields in the CS/PU of *X HEITTÄÄ HELMIÄ SIOILLE* |  $NP_{SUBJ}[N\{X\}]$   $V\{heittää\}$  'throw'  $NP_{OBJ}[N_{PL}\{helmi\}$  'pearl']  $PP_{ALL}[NP_{PL}\{sika\}$  'pig']] | lit. 'X throws pearls to pigs' are context-

dependent. They will vary depending on the particular #SITUATION#/PU, which the PU is used to conceptualize in discourse. The selected number of possible situations will be presented in Section 5.4.1 of Chapter 5 in connection with the variation analysis.

### 3.4.3.6.2 The modal tier

The topic of negation and negative modality in PUs that has been previously discussed in linguistic literature is reviewed in Section 2.5. In this section negative and modal features are located on the modal tier of the conceptual structure. Within the structure of PUs these features can be expressed explicitly, i.e. besides that they appear in the CS/PU, they can also be found in the CS with a counterpart in syntactic representation. For instance, the Finnish PU in (83) is a negative polarity unit that contains explicit negation in its form:

- (83) *X EI TULE HULLUA HURSKAAMMAKSI Y:STÄ*  
 lit. 'X **doesn't** become more devout than a madman from Y'  
 id. 'X **doesn't** understand Y'

On the other hand, there are PUs like those in (84) – (88). Dictionaries do not usually include explicit forms of negation and modality as a part of their morphosyntactic structure. Nevertheless, their CS/PU encodes a modal evaluation of the preformed action as inexpedient, unsuitable or pointless and therefore undesirable and inadvisable:

- (84) *HEITTÄÄ LAPSI PESUVEDEN MUKANA*  
 lit. 'to throw out the child with the bathing water'  
 id. 'to lose something valuable while getting rid of something unwanted'
- (85) *MAALATA PIRUJA SEINILLE*  
 lit. 'to paint devils on the walls'  
 id. 'to exaggerate potential risks'
- (86) *MENNÄ MERTA EDEMMÄS KALAAAN*  
 lit. 'to go fishing further than the sea'  
 id. 'to seek for something farther than is necessary'
- (87) *PITÄÄ KYNTTILÄÄ VAKAN ALLA*  
 lit. 'to keep one's candle under the bushel'  
 id. 'to keep one's talents a secret'
- (88) *HEITTÄÄ HELMIÄ SIOILLE*  
 lit. 'to throw pearls to pigs'  
 id. 'to cause a transfer of some entity, evaluated by the speaker as good, to a recipient, evaluated as inadequate in some way'

In the system of modality this evaluation falls within the scope of practical necessity, which means a necessity explained by ordinary reasoning: a situation is evaluated and on the basis of this evaluation a practical conclusion is drawn on whether a certain thing should or should not be done in order to achieve a wanted goal (Hakulinen et al. 2004: 1482). According to Laitinen (1993: 159), practical necessity is a semantic overlapper of dynamic, deontic and epistemic interpretation. In linguistic modal semantics it has been left somewhat aside, despite the fact that the semantics of practical necessity is quite close to the theory of pragmatic implicatures. Since this modal feature has not been previously described in Conceptual Semantics, I am introducing into my analysis a notation *PrNess*→*Neg*, where *PrNess* stands for ‘must’, ‘have to’, ‘ought to’, ‘should’ and *Neg* indicates negation.

A more detailed account of the negative modality of inexpediency in different constructions, as well as the morphosyntactic means of the *PrNess*→*Neg* features’ realization in discourse, will be presented later in Section 5.3.2.1 of Chapter 5 in the context of PU variation analysis. Since the current chapter is chiefly dedicated to the aspects of formal analysis, I will only present descriptions of two constructions from the *HELMIÄ SIOILLE* construction family given in (89) and (90) below. Negation is inherited into both constructions’ CS/PU from the negative imperative form of the PU’s original Biblical context<sup>127</sup> but it only appears in the phonological, morphological and syntactic structures (PS-MS-SS) of the latter construction:

- (89) *X HEITTÄÄ HELMIÄ SIOILLE* | *NP<sub>SUBJ</sub>[N{X}] V{heittää ‘throw’} NP<sub>OBJ</sub>[N<sub>PL</sub>{helmi ‘pearl’}] PP<sub>ALL</sub>[NP<sub>PL</sub>{sika ‘pig’}]]* | lit. ‘X throws pearls to pigs’
- (90) *EI HELMIÄ SIOILLE* | *NEG{ei ‘no’} NP<sub>OBJ</sub>[N<sub>PL</sub>{helmi ‘pearl’}] PP<sub>ALL</sub>[NP<sub>PL</sub>{sika ‘pig’}]]* | lit. ‘no pearls to pigs’, id. ‘one should not cause a transfer of some entity, evaluated by the speaker as good, to a recipient, evaluated as inadequate in some way’

Figure 70 demonstrates a formal description of the construction in (89). The *PrNess*→*Neg* elements appear only in the modal tier of the CS/PU. These modal features have no counterpart whatsoever in the CS or in the PS-MS-SS of this transitive verbal construction.

<sup>127</sup> *Älkää antako koirille sitä, mikä on pyhää, älkääkään heittääkö helmiänne sikojen eteen...* (1992, MATT 7:6) “Do not give what is holy to the dogs; nor cast your pearls before swine...” (NKJV, MATT 7:6)

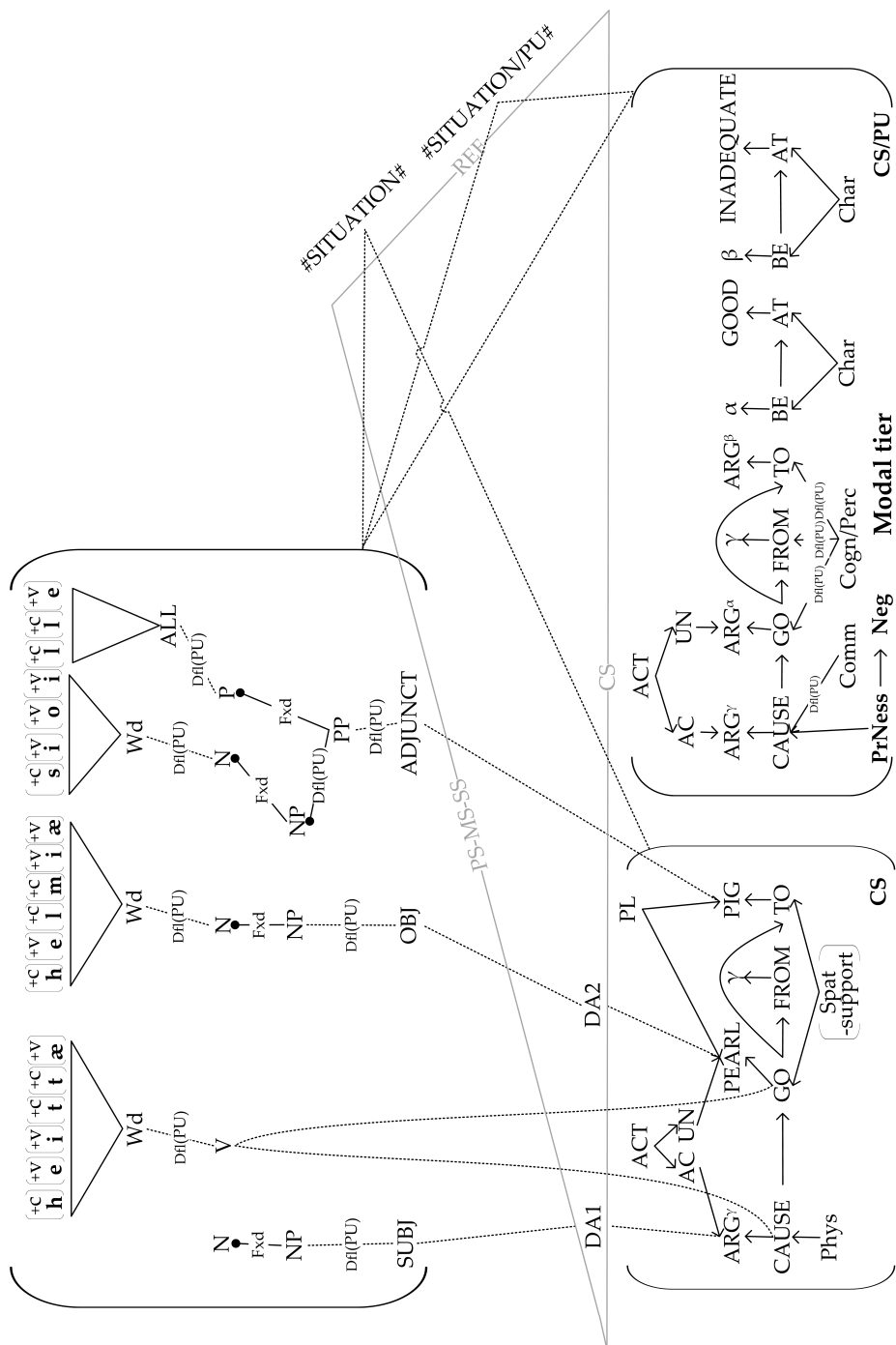


Figure 70  $PrNess \rightarrow Neg$  features in the modal tier of the transitive verbal construction *X HEITTÄÄ HELMIÄ SIOILLE* |  $NP_{SUBJ}[N\{X\}]$   $V\{heittää\}$  'throw'  $NP_{OBJ}[N_{PL}\{helmi\}$  'pearl']  $PP_{ALL}[NP_{PL}\{sika\}$  'pig']

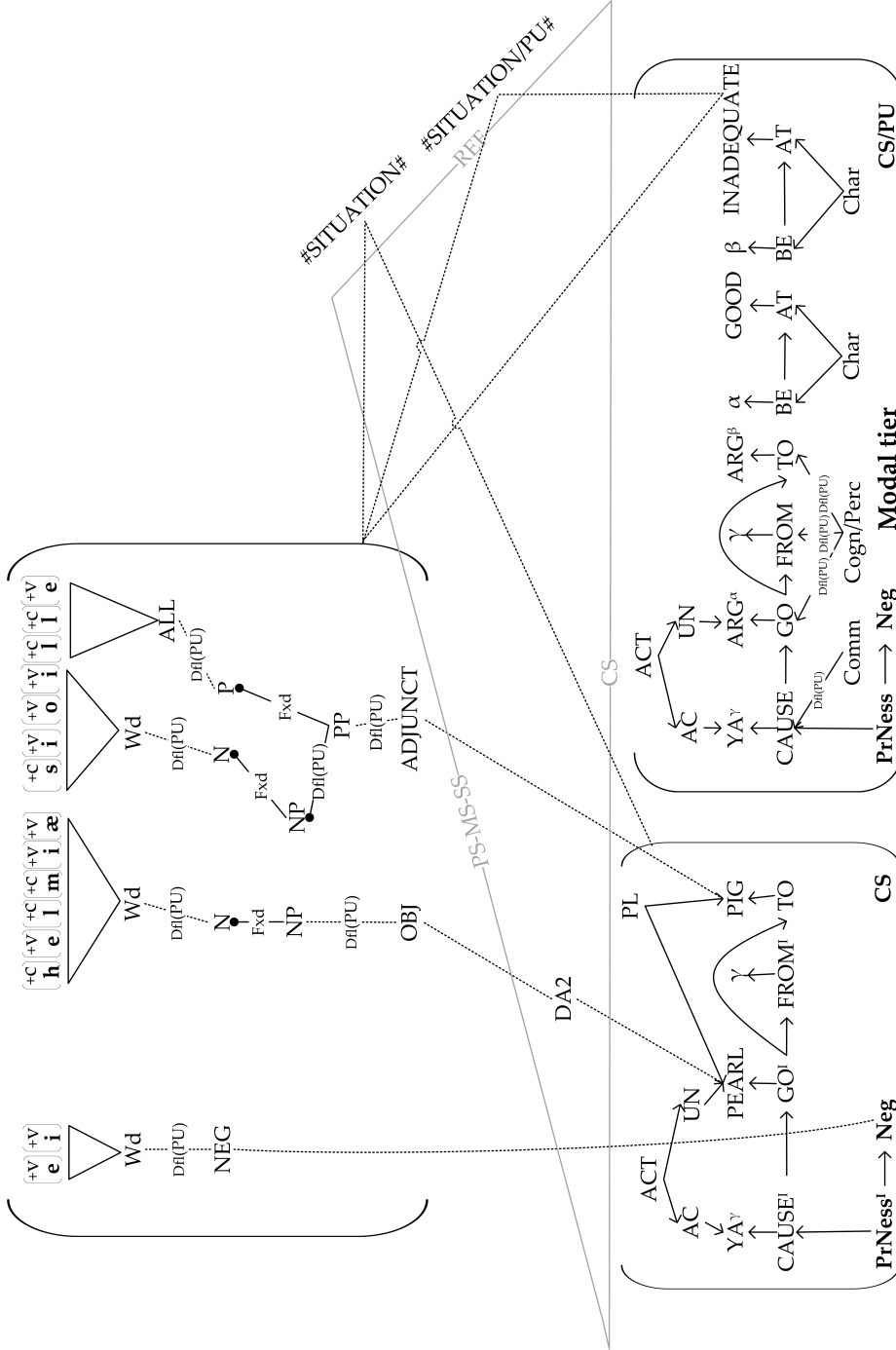


Figure 71 *PrNess*→*Neg* features in the modal tier of the verbless construction  
 EI HELMIÄ SIOILLE | NEG{ei 'no'} NP<sub>OBJ</sub>[N<sub>PL</sub>{helmi 'pearl'}] PP<sub>ALL</sub>[NP<sub>PL</sub>{sika  
 'pig'}]



In the verbless construction presented in (90) above, negation becomes overtly manifested, i.e. one can find its counterpart in syntactic representation. The construction's formal description is shown in Figure 71 above. As one can see, the NEG particle<sup>128</sup> in syntax corresponds to *ei* in phonology and to the Neg feature in the CS. Besides the negative particle, the PS-MS-SS of this construction does not feature any elements that could be connected to the modal PrNess feature in the CS. For this reason, it is marked as implicit with a superscript *I*<sup>129</sup>.

### 3.5 Summary

In the beginning of this chapter (Sections 3.1, 3.2 and 3.3) I briefly looked at some theoretical assumptions, methodological guidelines and formal apparatus of the Conceptual Semantics and the Tiernet model. As Nikanne (2005a, 2008a) points out, Conceptual Semantics pursues to be a formal approach with the analytical organization of the model. Since the principles of representational modules are kept as simple as possible, the linking principles between representations become an increasingly important part of the model. Linking does not need to be a one-to-one relation, but it is important to distinguish between regular and irregular linking patterns. Nikanne (2008a) formulates the latter guideline as Regularities before irregularities, i.e. even though the irregularities are important, the possibility of referring to regularities has to be checked first. In the same vein with Nikanne's (2005a) view of constructions, it has been argued that PUs have to be treated as linking devices that license irregular linking patterns between the tiers of syntactic and conceptual representation. It has been argued that the formal distinction has to be made between at least three types of linking relations: constituency, selection (dependency) and correspondence (Section 3.2.1).

Section 3.4 is dedicated to theoretical and methodological aspects of formal description of PUs in the Tiernet model. I presented different parts of the network structure relevant for the PU analysis and the linking

---

<sup>128</sup> The negative *ei* can function either as an indeclinable particle or as an auxiliary verb (it would be more correct, though, to call it a functional head in a finite clause), when it agrees with the grammatical subject (*e-n*, *e-t*, *e-i*, *e-mme*, *e-tte*, *e-ivät*) while the finite verb appears in its negative form.

<sup>129</sup> Note that there are other implicit elements in the CS of this construction: since there is no verb, the thematic functions CAUSE, GO and FROM are also marked as implicit.

between these parts. In connection with the description of a PU's lexical constituents' phonological and morphological form, a possibility has been discussed for each linking type to be given different values, indicating the strength of a particular connection, e.g., fixed, default, non-default and unmarked linking. The totality of default linking for a given PU results in its default form (Section 3.4.2). In contrast to fixed linking, default value allows for variation. In Section 3.4.1.3 it was demonstrated that the lexical constituents' conceptual structure could include non-linguistic fragments of representation in the form of qualia structure and evaluation. Exemplified by the evaluative component of the qualia structure, it has been shown that these fragments can form salient motivating links to the CS/PU (i.e. conceptual structure licensed by the PU).

The problems of literal meaning and analyzability were discussed in Sections 3.4.3.1-3.4.3.3. The notion of literal meaning has been abandoned as vague and incompatible with the view that conceptual structure is a single level of semantic representation. Instead, I used the notion of CS licensed by the regular linking rules, as opposed to the CS/PU licensed by the PU. I argued for the necessity to include both conceptual structures into the formal description of ambiguous PUs. The CS/PU cannot be directly linked to syntax; instead this linking is assumed to be done via the referential tier. Thus, the analyzability of PUs is understood as the PU referentiality of its syntactic constituents (Sections 3.4.3.3 - 3.4.3.4). Reference in PUs does not follow the general Referentiality Principle. Some principles of PU referentiality were discussed in Section 3.4.3.4. The referential tier must be regarded as an important part of the formal representation of PUs, since it explicates referential relations that underlie establishing the contextualized meaning of these expressions. In semantically ambiguous PUs (i.e. PUs with unspecified thematic arguments in the CS/PU) like *HEITTÄÄ HELMIÄ SIOILLE* 'throw pearls to pigs', disambiguation demands for the referential indeterminacies to be resolved (Section 3.4.3.5). Finally, in Section 3.4.3.6, I looked at the semantic fields and the modal tier as a part of PUs' formal description. A discrepancy between semantic fields in the CS and the CS/PU was demonstrated in a couple of PUs. In the modal tier, the notation *PrNess*→*Neg* was introduced to encode a modal evaluation of inexpediency.

## 4. Critical overview of previous approaches to PU variation

The irregular lexicogrammatical behaviour of PUs is assumed to result in certain restrictions in the choice of their morphosyntactic and lexical forms. The general term for this feature is frozenness (Fraser 1970), although some scholars make a further distinction between morphosyntactic fixedness, which corresponds to restricted structural variability, and restricted collocability, which corresponds to lexical invariability (e.g. Langlotz 2006: 4). Within the framework of transformational grammar, the inability to undergo all grammatically possible transformations is referred to as transformational deficiency (Weinreich 1969: 47).

PUs have always been described as more or less stable units with reference to their form and meaning. An important role in their recognition and interpretation has been assigned to their lexicogrammatical and semantic structure. Restricted morphosyntactic and lexical variability has been used as one of the main definitory parameters that distinguish PUs from free syntax. Nevertheless, lexicogrammatical variability of PUs is a recognized empirical fact. Recent corpus studies of PUs (e.g. Moon 1998, Fellbaum 2007, Sköldberg 2004, Langlotz 2006, Heinonen 2007) tend to show that their structure can actually undergo considerable modifications. Although stability of form and meaning is regarded as an important characteristic feature of PUs, it is important to keep in mind that their stability is indeed relative and their form is far more unstable than is often assumed.

The tendency to ignore or diminish the general character of this phenomenon is reflected in the terminology used by researchers who emphasize the singleness and uniqueness of variants by using such terms as occasional transformations, occasional derivation, occasional variant, instantial stylistic use or author's individual modifications. The latter term, which can be found especially in Russian phraseological theory, owes its coinage to the fact that before large computer corpora became available, the majority of empirical research in this field was performed on the basis of data obtained from fictional literature; in many cases examples of alternated PUs were collected for the purpose of a particular research from the

works of a single author (e.g. Boychenko 1993, Podgurska 1990). In this context it seems more natural for a scholar to label modifications as author's individual and describe them mainly in stylistic terms. As long as the language use of ordinary people was concerned, modifications were referred to as something undesirable, incorrect or even abnormal. This attitude became manifested in the terminology like deformation, distortion, manipulation etc (for more discussion see Naciscione 2001: 7).

This chapter presents a critical overview of some previous approaches to (mainly formal) variation in PUs. By formal variation here I mean all kinds of discrepancies that can exist between a default morphosyntactic and phonological form of a target unit, on one hand, and, on the other hand, actual tokens of its occurrence in data. Section 4.1 is devoted to a brief discussion of some syntactic and semantic accounts of frozenness. Syntactic view is represented by frozenness hierarchies (Fraser 1970, van Gestel 1995, Tronenko 2003). Alternative semantic models (Chafe 1968, Newmeyer 1974, Burger 1973, Nunberg, Sag & Wasow 1994) try to correlate syntactic flexibility with semantic analyzability. As will be shown, both approaches have problems, since they are based on intuitive judgments, do not take an idiom's context into account and can often be contradicted by empirical evidence. Section 4.2 aims at a closer look at different categories of variation and principles of categorization. What is the difference between variation and transformations? Can one divide variation into purely semantic or purely syntactic? Is it methodologically sound to label a variant as usual or occasional? What is systematic variation and how does it differ from idiomatic wordplay (Langlotz 2006)? The methodological efficiency of describing variation in terms of categorization of its different types can often be questioned.

Findings from the recent corpus-based studies of phraseological units' behaviour show that PU variation is not as occasional or individual as one might think. The evidence for variation is in fact so convincing that it even questions the notion of the canonical form and suggests that it should be superseded and other models should be developed instead (Moon 1998:121, Pulman 1993). This view contradicts the common practice to consider variants as being opposed to the PU's base (canonical, neutral) form and core (neutral, standard) use. The former refers to "a decontextualized unit in the system of language, an

abstraction, a context-independent default-structure” (Langlotz 2006:176). The latter denotes the most common form in which a PU appears in discourse, “the standard usage of an idiom according to its base-form. Uses that deviate from this standard are defined as variations or alterations” (Langlotz 2006: 176). However, a closer look at the problem reveals a heuristic nature of the base form approach, which will be discussed in Section 4.4 of this chapter. I will argue that the traditional view of variation as manipulations with a postulated base form should be abandoned.

#### **4.1 Problems in some previous syntactic and semantic accounts of PU frozenness and variability**

The treatment of PUs within the framework of transformational grammar has been marked by several attempts to find regularities in their syntactic behaviour and to systematize their recalcitrance to undergo certain syntactic transformations. The transformational hierarchy for English idioms was first suggested by Fraser (1970), who observed that idioms with the same syntactic structures do not permit the same set of transformations. According to Fraser (1970: 23), idioms can range from “completely frozen” to “fairly amenable to transformational operations”, i.e. frozenness is gradable. The ultimate poles are represented by levels L0 (completely frozen) and L6 (unrestricted). All variation happens on levels L1 to L5: Adjunction, Insertion, Permutation, Extraction and Reconstitution. Fraser (1970) points out that L6 is not applicable to idioms, i.e. operations like clefting, relative clause formation, conjunction reduction, gapping and pronominalization are not possible in idioms. According to Fraser (1970), the most significant feature of this hierarchy is that any idiom marked as belonging to one level is automatically marked as belonging to any lower level. The first problem is that Fraser’s (1970) claims are supported only by self-constructed sentences, i.e. he relies solely on his own intuition when assigning a level of hierarchy to an idiom. He admits the fact that the place of a particular idiom in the hierarchy may differ among speakers, but he emphasizes that the system as a whole is nevertheless valid for each speaker. However, Dong’s (1971) and Newmeyer’s (1974) criticisms of Fraser’s (1970) model include examples that contradict with his claim of its postulated systematic nature, e.g. Newmeyer (1974) points out that *cast pearls before swine* can be

passivized but does not allow a lower level transformation of adverb preposing. There are also psycholinguistic studies (Nenonen 2002) that point to the absence of a corresponding hierarchy in Finnish, i.e. Fraser's (1970) model is not universal.

There have been several attempts to account for hierarchical order of degrees of idiomatic fixation in terms of X-bar projection levels (van Gestel 1995, Tronenko 2003). In his study of Dutch prepositional idioms, van Gestel (1995: 80) argues that "degrees of idiomatization are level-determined: The higher the X-bar level, the higher the degree of fixation". Tronenko (2003) applies this view to Russian phrasal idioms and suggests a frozenness hierarchy with respect to the constraints that their verb constituent is subject to. Tronenko's (2003) claim is that idioms may have specified functional Asp, Tense and Agr features of V (Figure 72 below). Tronenko (2003) makes a strong claim that elimination of the restrictions of any of the parameters would lead to the loss of idiomaticity. Her hierarchy contains the following levels:

*V-idioms and V'-idioms* with no categorical restrictions > *Asp'-idioms* with aspectual restrictions > *T'-idioms* with temporal restrictions > *AgrsP-idioms* with subject agreement restrictions.

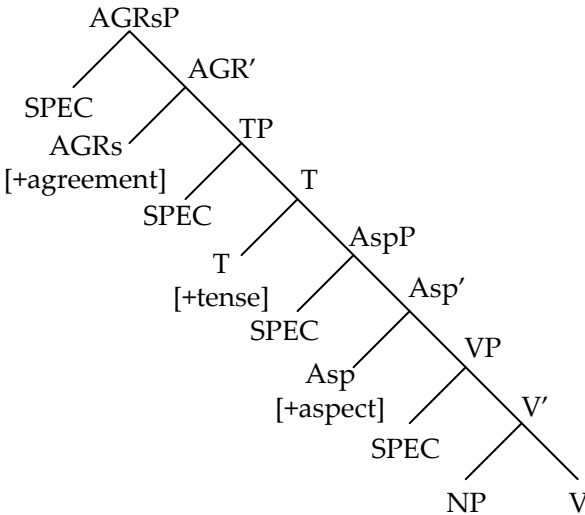


Figure 72 X-bar projection levels of verbal categorical properties.

Here, just like in Fraser's (1970) model, higher level constraints should be valid for lower levels, i.e. phrasal idioms restricted in terms of their subject agreement are also restricted in aspect and tense. Tronenko (2003)

remarks that along with other categorical properties, the mood of the idiomatic verb can also be restricted, but chooses not to include this category into her hierarchy “for the sake of clarity”. Concerning the fact that different categories of mood in Russian interact with tense and person agreement in a different way, it could indeed introduce some complications into the postulated system: e.g. verbs in the imperative and subjunctive mood do not express temporal differentiations; imperative mood forms are restricted to second person and forms of joint action. As for Asp’ idioms, not all verbs in Russian are able to form aspectual pairs, i.e. the restriction could be imparted by the verb itself and not by the idiom.

Tronenko’s (2003) analysis is based only on intuitive and introspective examples and acceptability judgments that can be easily contradicted. Of course, one could resort to Fraser’s (1970) argument and claim that although the position of an idiom in the hierarchy can vary in different speakers’ idiolects, the principle still remains the same, i.e. higher level constraints are applicable in any case. However, there are examples that contradict the systematicity of Tronenko’s (2003) hierarchy. For instance, she places the Russian PU *ДАЙ БОГ НОГИ* [*daj bog nógi*] lit. ‘give<sup>PERF IMPV SG2</sup> God legs’, id. ‘run off quickly’ on the highest level of frozenness, which implies that its verb is constrained both aspectually and temporally and is restricted to a “bound paradigmatic form of person-gender-number agreement” (Tronenko 2003: 143). The verb *дать* [*dat’*] ‘to give’ here is bound to the imperative mood second person singular form. Given that imperative in Russian lacks temporal forms, the verb can hardly be regarded as temporally restricted by the idiom. But most importantly, the idiom has a conventional alternative aspectual form *ДАВАЙ БОГ НОГИ* [*daváj bog nógi*] (listed in several dictionaries, e.g. Lubensky 1997: 20) where the verb *давать* [*davát’*] ‘to give’ is the imperfective aspectual pair of the perfective *дать* [*dat’*]:

- (91) Успел на обратный автобус заскочить и **давай бог ноги** из этого царства навозной вони.<sup>130</sup>

[*Uspél na poslédnij avtóbús zaskočít’ I daváj bog nógi iz éтого cárstva navóznoj vóni.*]

lit. ‘(I) was in time to catch the return bus and **give<sup>IMPERF IMPV SG2</sup> god legs** from this kingdom of dung stench.’

<sup>130</sup> GG: fido7.ru.pol.opposition/msg/98ef0e0c527a06d9, Oct 20 2006

id. '(I) was in time to catch the return bus and quickly ran away from this kingdom of dung stench.'

Here we have a phrasal idiom with a verb in a bound person-number form but appearing both in imperfective and perfective aspectual forms – something, which, according to Tronenko's (2003) hierarchy, would be non-idiomatic.

Syntax-based approaches to idiom variation have also been criticized for a lack of explanatory power, since they describe transformational deficiency as something given, without trying to explain the reason why certain idioms are more frozen than others (Burger 1973: 70, Schenk 1995: 255, Langlotz 2006: 22). However, this does not necessarily mean that alternative semantic models (Chafe 1968, Newmeyer 1974, Burger 1973, Nunberg, Sag & Wasow 1994) are capable of explaining the syntactic behaviour of idioms in purely semantic terms. For instance, in their pursuit to correlate meaning with syntactic flexibility Chafe (1968), Newmeyer (1974) and Burger (1973) claim that passivization cannot be applied to the notorious idiom *KICK THE BUCKET* since it describes an intransitive process of 'dying' and intransitive verbs are not open to passive voice. In a similar way Chafe (1968: 122) appeals to semantic analyzability when he states that the NP *the bucket* cannot be modified because it does not correspond to any constituent of the idiomatic meaning. In the light of such a semantic approach it would be difficult to explain the existence of examples like (92) – (99) below, where this idiom is passivized and (95) – (101), where its constituent *the bucket* is modified. Moreover, according to Nunberg, Sag & Wasow (1994: 500-503) only the literal reading would be possible for the pluralization transformation to *buckets* as in (101) below, which is apparently not the case.

- (92) *I reckon one aspect of it \*truly\* being someone's time is when the deal is done and **the bucket kicked**.*<sup>131</sup>
- (93) *When **the bucket is kicked**, the body battery is turned off. No brain waves, electrical activity, just a slab of beef (unless you're hindu).*<sup>132</sup>
- (94) *But I always figure, if I live what I consider to be a good life, then the problem can safely be put off till after **the bucket is kicked**; God and I can argue about the nitpicky details then.*<sup>133</sup>

---

<sup>131</sup> GG: alt.religion.shamanism/msg/52de20f8d51a2228, Jan 21 2003 (a message from the discussion thread *death (sort of OT)*)

<sup>132</sup> GG: alt.agnosticism/msg/73e4f5143a530dfd, Feb 17 1999

<sup>133</sup> GG: rec.arts.sf.written.robert-jordan/msg/014c9d2a17549e4b, Mar 7 2000



- (95) *Oh gimme a break, he was just another stroppy student until **his bucket was kicked**.. Talk about posthumous sainthood.*<sup>134</sup>
- (96) *Jeez, **another bucket kicked**.*<sup>135</sup>
- (97) *geezzzzz... if you decided to **have yer bucket kicked**, forget yer ecological principles for once, go to the gas station and get one or two tanks full of leaded fuel... once they're "blown" out the exhaust, do the smokey garage experience again... should be more CO than you can handle..*<sup>136</sup>
- (98) *You shouldnt need an ISA so your ISA contribution limits can be saved for your own use but mark the contribution to the UT as being for the account of your son, and put something else in writing to the same effect, so that if **your bucket is kicked** before he needs the dosh then the funds shouldnt form part of your estate.*<sup>137</sup>
- (99) *If you expect to go to heaven when **your bucket is kicked**, then I can expect to receive undoubtable proof direct from the source that there even IS a heaven, before I do anything towards going there.*<sup>138</sup>
- (100) *Poor old Spike has **kicked his bucket**.... age 84*<sup>139</sup>
- (101) *One of the great ironies of the 20th Century: Jews such as Julius and Ethel Rosenberg, who tirelessly worked for Stalin, would have met the same fate in their beloved Workers' Paradise that they eventually did at Sing Sing. Only in Siberia, they would have been a \*lot\* thinner when they finally **kicked their buckets**...*<sup>140</sup>

The above-listed examples show that even semantically non-decomposable PUs like *KICK THE BUCKET* allow some transformational productivity, which means that the syntactic variability and transformational deficiencies of non-compositional PUs cannot be predicted solely on the basis of their semantics<sup>141</sup>. In a similar way, the fact that an idiom's literal counterpart is passivizable or the presence of a passive-governing predicate in the idiomatic meaning does not

<sup>134</sup> GG: za.politics/msg/a8bb7c918b0993f6, Feb 13 2000

<sup>135</sup> GG: alt.fan.frank-zappa/msg/6cd456cbfa9c4d65, Oct 12 1999 (a message from the discussion thread *If Wilt The Stilt and FZ be jammin' etc...* posted on the date of Wilt Chamberlain's death)

<sup>136</sup> GG: sci.chem/msg/6af99d25c0b36029, Jul 9 2003

<sup>137</sup> GG: uk.finance/msg/8a195aa912ce0ac6, Jun 19 2001

<sup>138</sup> GG: talk.atheism/msg/05eef14df1f81735, Aug 25 1999

<sup>139</sup> GG: free.uk.btinternet.chatter.refugees/msg/49d51acecc485c37, Feb 27 2002

<sup>140</sup> GG: can.politics/msg/f282e45b6548bb53, Mar 13 2000

<sup>141</sup> It is worthwhile mentioning that traditional corpora are unable to provide with such examples, e.g. Moon (1998: 109) notes that there were no passives among the 42 tokens of *KICK THE BUCKET* in the Bank of English corpus, but she herself admits that "the absence of passive forms is negative evidence that proves nothing either way" (ibid. 105).

necessarily increase the idiom's passivization variation potential. This argument is supported empirically by Moon (1998: 109), who observed a strong fossilization of *SPILL THE BEANS* in active voice (only 4 of 198 tokens were passive), despite its potential to be passivized. In the same vein, Abeillé (1995: 24) remarks that there is no clear correlation between syntactic regularities and semantic analyzability.

It is true that acceptability of a particular PU form can vary from speaker to speaker. But it is also true that PUs change their form only when they are used in actual discourse. Therefore any sophisticated reasoning about non-availability of idiomatic interpretation, in case a certain alternation is applied to a PU, can be proved wrong if it does not take into consideration the simple fact that PUs never function outside context. It is obvious that the role of context in processing formally and semantically alternated PUs should not be underestimated. The same point of view is shared by Burger et al. (1982: 68), who notice that there are hardly any types of variation of an idiomatic expression, which would not be feasible given an appropriate context. Pulman (1993) makes a similar point by saying that it seems legitimate to regard all idioms as being able in principle to occur in any syntactic configuration, but in his opinion syntactic forms would "sound odd if there is no context in which the way they present information is plausible". Thus, focusing on *the bucket* in *KICK THE BUCKET* is "odd", if it does not correspond to any discourse entity, which it makes sense to emphasize.

## 4.2 Categorical models of variation – oppositions and classifications

The majority of previous studies in PU variation were aimed at distinguishing and describing different types of such. Accordingly, every author engaged in the description of a variety of forms that occurring in his/her data would naturally present his own categorization of these phenomena. The explanation of differences can be found in the nature of classical categorization: categories do not represent something which is objectively present in the data, but are created for a specific purpose – in order to group examples according to their similar properties. Different categorizations are achieved by emphasizing different properties of examined entities. In this section I will discuss and compare categorization principles of PU modification that occur in linguistic

literature. My claim is that most of them fail to fulfil the requirements of keeping categories clearly defined, mutually exclusive and collectively exhaustive. The complexity of language phenomena simply does not allow fitting them into the Procrustean bed of taxonomic approach.

#### **4.2.1 Variation vs. transformations**

One cannot but notice the concurrent use of the terms transformations and variation in literature on idioms. Quite often they are used synonymously with either one of them appearing as a major hypernym that covers all kinds of performed changes, although in some cases they are regarded as separate categories (e.g. Moon 1998). As will be shown below, there is a great matter of inconsequence in the use of these terms and their further subcategorization.

Established within the framework of transformational grammar, the term transformation was originally used to denote a formal linguistic operation on the constituents of a deep structure which converts into a surface structure. It is worth mentioning that the amount of basic operations (called elementary transformations), their definitions and status, as well as names given to them, vary within the theory. However, a full and adequate discussion of this issue is beyond the scope of this thesis. Crystal (1997: 394–395) mentions movement (reordering or permutation), adjunction, insertion and deletion as basic operations. Chomsky (1965: 144) himself advocated elimination of permutations from the set of elementary transformations for the sake of simplification. His base set of elementary transformations included adjunction, substitution and deletion. In Chomsky's (1965) terms specific grammatical transformations, like nominalization, passivization etc., are sequences of elementary transformations. At the same time there have been attempts to use elementary transformations for categorization purposes. Jacobs et al. (1968: 26) admit that within elementary transformation types there are probably subtypes. Chomsky (1965: 143) also remarks that transformations may refer to specified syntactic features as if they were categories.

Fraser's (1970) Frozenness Hierarchy could be regarded as an attempt to categorize morphosyntactic changes in idioms in terms of elementary transformations, since he claims that the types of operations he makes use of (adjunction, insertion, permutation, extraction, reconstitution) are not defined in terms of any particular grammatical transformations. He calls

them operations on P-markers, which specific transformations make use of to map a deep structure to a surface structure P-marker. This definition in fact corresponds to the notion on elementary transformations although Fraser (1970: 37) himself states the opposite. It is still true that Fraser's (1970) set of operations differs from Chomsky's (1965) base set, except for adjunction, which can be found in both. Nevertheless, the status of Fraser's (1970) operations remains unclear.

Despite the fact that Fernando's (1996) functional treatment of idioms lies outside the framework of transformational grammar, she still chooses to operate with the base set of elementary transformations while categorizing various ways in which idioms can be changed. Her classification includes replacements (substitutions), additions, permutations and deletions. Apparently, this kind of classification makes no difference if the described phenomena occur on the morphosyntactic or lexical level. Its four classes are therefore aimed to represent all possible types of qualitative and quantitative changes in form. Consequently, each category involves phenomena of very different kinds. For instance, replacements (substitutions) embrace inflectional changes (variation in number and tense), replacement of structural words like articles by another or zero and substitution of content lexical constituents – both fixed ones and so-called built-in variables<sup>142</sup>. Permutations include passivization, particle shift, conversion of a verb + predicate into a nominal (Fernando calls them compressions) and reversal of subject and object. The apparent problem with this kind of classification lies in the fact that a specific grammatical transformation by definition cannot be described under a single elementary transformation category, since the former involves a sequence of several operations. For instance, several examples presented by the author as deletion also include addition and replacement, passivization involves both addition and permutation, replacement of structural words by zero could also be regarded as deletion, etc.

Moon (1998) distinguishes between transformations and variation as two different categories, but the reason for this distinction is not quite clear. Transformations are discussed together with such issues related to lexical and grammatical form of idioms as grammatical types and

---

<sup>142</sup> These are called open (regular) slots in Moon's (1998: 98) terminology.

structures, inflectability and regular slots<sup>143</sup>. Moon's transformation subcategories involve specific grammatical transformations: passivization, nonfinite uses, embedding, pronominalization, nominalization, etc. On the other hand, variation includes a subcategory of lexical variation, which could have been analyzed in terms of elementary transformations as well. For instance, constituent variation (e.g. *KEEP/JUGGLE THE BALLS IN THE AIR*) could be treated as substitution, specificity and amplification (e.g. *CUT THE CORD* vs. *CUT THE UMBILICAL CORD*) as insertion, truncation (e.g. *MAKE HAY WHILE THE SUN SHINES* vs. *MAKE HAY*) as deletion, and reversals (e.g. *DAY AND NIGHT* vs. *NIGHT AND DAY*) as permutations. Examples of the same construction can be found appearing in transformation and variation categories simultaneously: e.g. truncation of a verbal idiom to a nominal phrase is claimed to be a specific form of nominalization. But while nominalization is listed among transformations, truncation is considered to be a subtype of lexical variation.

Finally, there are authors who do not use the term transformations at all and speak only about variation. Heinonen (2007) distinguishes three kinds of variation in VP idioms: grammatical, lexical and constructional. However, these categories are far from being mutually exclusive. As a matter of fact they constantly interact: lexical variation is involved both in grammatical operations and in constructional variation; grammatical operations, like passivization, nominalization etc., create new phrase patterns and therefore result in new constructions.

#### **4.2.2 Some major categories of variation and their definitional criteria**

In this section I will discuss some general principles of variation categorization. It will be argued that there is no such category as purely formal variation and that the distinction between usual and occasional variation is subjective since it depends to a large extent on the size and make-up of the corpus available to the researcher. There are practically no valid criteria for the definition of occasional variation. A large part of this section is devoted to an overview and criticism of the notions of

---

<sup>143</sup> Regular slots are treated as transformations by Fernando but fall outside this category in Moon's (1998) classification.

systematic variation and idiomatic wordplay as they appear in Langlotz (2006).

#### 4.2.2.1 Formal vs. semantic variation

Dividing variation into purely semantic (i.e. variation in an idiom's meaning), on the one hand, and purely formal (i.e. variation in an idiom's form), on the other hand<sup>144</sup> (e.g. Sköldberg 2004, Langlotz 2006), neglects the fact that changes in form would inevitably entail changes in meaning. For example, the meaning adaptations subtype of semantic variation which can be found in a technical classification of idiom variation proposed by Langlotz (2006: 179), involves phenomena of formal variation as well: intensifications in (102) below<sup>145</sup> and specifications in (103) involve pre- and postmodification, while perspectivization in (104) implies a change between an intransitive motion (a) and a caused motion constructions (b) (Langlotz 2006: 181-182):

(102) [...] *this time overturning **the most ponderous applecart of all*** (ibid. 181).

(103) *Preaching hatred, as people like Ian Paisley do, only fans the flames **of vengeance*** (ibid. 182).

(104) (a) *the curtain **comes down** on sth.*

(b) ***bring** the curtain **down** on sth.* (ibid. 182)

Melerovič & Mokienko (2001) come up with the following solution: they do not have a category for purely formal variation. While the first higher level category in their classification is supposed to include solely semantic modifications, which do not affect the lexicogrammatical structure of idioms, the second category of modifications is called structural-semantic, thus emphasizing that changes in meaning are connected with alternations in lexicogrammar. In practice, however, the authors fail to keep purely semantic and structural-semantic changes apart from each other when it comes to actual examples<sup>146</sup>. Numerous examples listed under the category of semantic transformations, in

---

<sup>144</sup> Terminology can vary, e.g. Sköldberg (2004) labels variation in meaning *innehållsmässig* (Swe 'content-wise') and variation in form – *uttrycksmässig* (Swe 'expression-wise') while Melerovič & Mokienko (2001) use the terms semantic and structural to describe the same phenomena.

<sup>145</sup> This example involves another kind of formal variation, namely lexical substitution of the verb *upset* with *overturn*.

<sup>146</sup> Melerovič & Mokienko (2001) classification of variation appears in the preface to their dictionary of phraseological units in Russian speech, whereas examples of variation are given in the dictionary entries.

fact contain structural changes as well, e.g. the Russian PU *НЕОСТАВЛЕНИЕ КАМНЯ НА КАМНЕ* [*neostavlenie kamnja na kamne*] lit. 'not leaving one stone upon another', id. 'total destruction', presented by Melerovič & Mokienko (2001: 270) as belonging to the subcategory literalization, involves nominalization of the verbal idiom *КАМНЯ НА КАМНЕ НЕ ОСТАВИТЬ* [*kamnja na kamne ne ostavit'*] lit. 'not to leave one stone upon another', id. 'to destroy totally'.

#### 4.2.2.2 Usual vs. occasional variation

The terms usual variant and occasional variant are commonly used to distinguish between frequently recurring vs. non-recurrent idiom variants in usage-based studies of idiom variation (e.g. Burger 1998; Burger et al. 1982; Melerovič & Mokienko 2001; Langlotz 2006). According to Langlotz (2006: 199), the terms are supposed to "capture differences in the commonness and institutionalisation of idiom variants": usual variants are "recurrent and institutionalized", occasional are "transitory and restricted" to one specific usage-event. Can usual be equalled to institutionalized? Apparently not, at least for Langlotz (2006: 177), who makes a distinction between an institutionalized variant as "an idiom variant that has become institutionalised" and a usual variant as "a frequently recurring idiom variant". Of course, one can argue that institutionalization as a sociolinguistic concept can be defined by more parameters than just a mere frequency of reproduction. But on the other hand, what other means of proof that a certain pattern is institutionalized, apart from frequencies observed in the data, are available to a scholar who studies written corpora and does not attempt to measure institutionalization by sociolinguistic tests?

Another problem with these terms is that they are supposed to reflect the conventionality within a given speech community. In practice, they only really describe the situation in data available to a researcher. Therefore, when labelling a variant as occasional, it is important to remember that it might be non-recurrent only in this particular data and not to make further generalizations. It is obvious that even large computer corpora have limitations. For instance, Langlotz (2006: 226–227) remarks that since the approximate range of idiom tokens per type in the 100 million-word British National Corpus<sup>147</sup> is only 10–50 and idiom

<sup>147</sup> The British National Corpus (BNC) is a 100 million-word collection of samples of written and spoken language from a wide range of sources, designed to represent a wide

variants are expected to appear even more infrequently, stable quantitative predictions about idiom variability cannot be derived on the basis of corpus analysis. Considering this observation, as well as the fact that Langlotz himself does not use the corpus for quantitative analysis, the purpose of the classificatory distinction between usual and occasional is unclear. On the contrary, Moon (1998), who presents a quantitative corpus analysis of idiom usage and can therefore talk about frequencies of variation with actual numbers at hand, does not include the usual/occasional categories in her classification.

There are also many other questions that remain unanswered. What are the quantitative criteria for a variant to be occasional vs. usual? If occasional variants are indeed restricted to only one specific usage event (Langlotz 2006: 199), how shall one treat those variants that occur two or more times in the data? For instance, when Langlotz (2006: 205) finds in the BNC two examples where the verb *upset* in the English idiom *UPSET THE APPLE CART* is substituted with the verb *rock*, he notes that although the variant is not listed in dictionaries the recurrence of substitution “reflects some degree of commonness” and therefore “cannot be described unequivocally as an occasional form”. If one presumes that usual and occasional are not clear-cut categories and considers them as two poles on a continuum of commonness, then how should one proceed in defining the transitory zone between them? And finally, since institutionalization is usually considered as one of the definitory features of idioms, what prevents one from treating a usual (or institutionalized) variant as an idiomatic unit in its own right?

Apart from being frequent and institutionalized, usual variation is claimed to be systematic:

The notion of usual variation overlaps with the notion of systematic variation to the extent that it involves codified systematic departures from an idiom’s base-form (Langlotz 2006: 199).

Systematic variation extends to the category of occasional variation, which also includes non-systematic wordplay. The categories of systematic variation and wordplay together with their definitional criteria will be discussed in detail in the next section.



#### 4.2.2.3 Systematic variation vs. wordplay

The notion of systematic variation can be found in e.g. Moon (1998) and Langlotz (2006), where it is defined in quite a different manner. Moon (1998: 139) uses the term systematic variations to describe “deeper grammatical systems and relationships or concepts”<sup>148</sup> than “simple transformations and grammatical operations”, such as passivization, polarity, nominalization etc., which she discusses as “part of the routine morphological behaviour” of idioms. Systematic variations “display some sort of regularity”, “they may be predicted to occur in text, although this does not necessarily mean that they do occur”. This category includes cases of alternative constructions, expressing variation in the notion of possession, causative and resultative structures, aspect, reciprocity etc. Heinonen (2007: 147) calls this type of variation constructional.

The notion of systematic variation plays a crucial role in Langlotz’s (2006) cognitive model of idiom-variation:

The aim of this study is to carve out a cognitive idiom-variation grammar that can account for the **systematic variability** of idiomatic constructions. Thus, we are now forced to find criteria to define the notion of **systematic variability**. It is of primary interest to chart the fuzzy area between **systematic and non-systematic idiom-variation** (Langlotz 2006: 194; boldface added by me).

Langlotz (2006: 8, 177) defines a systematic variant as “an idiom variant that can be described as regular and grammatical”, involving “a systematic application of regular grammatical processes”. On the basis of this definition it would be logical to expect the opposite category to include cases of irregular and ungrammatical forms. Instead Langlotz (2006) makes a distinction between systematic variation and (non-systematic) wordplay. But, obviously, wordplay cannot be described as something ungrammatical, so there has to be some other criteria underlying this distinction. Indeed, Langlotz’s (2006) notion of systematic variation includes more parameters than the above-mentioned definition suggests. These parameters are motivation by conceptual patterns of semantic extension, weakly implicated contextual effects and idiom-variation principles which reflect the communicative

---

<sup>148</sup> Moon (1998) does not really explain these terms, so it is not quite clear what exactly the epithet “deeper” refers to.

function of variants. In the following sections I will analyze these criteria of systematicity and demonstrate that they to a large extent rely on subjective factors and poorly defined concepts.

#### 4.2.2.3.1 Motivation by conceptual metaphors as a criterion of systematicity of variation

Langlotz (2006: 193–194, boldface added by me) claims that systematic variation of idioms depends on their motivation:

To the extent that the association between the literal and the idiomatic scene can be motivated and analysed on the basis of underlying conceptual patterns of semantic extension, the cognitive micro-model can be **systematically manipulated**. In the absence of motivating conceptual bases, an idiom's potential for **systematic variation** is restricted.

Motivation, in Langlotz's (2006: 45) terms, refers to a speaker's ability to "make sense" of an idiomatic expression, i.e. to understand why it has the idiomatic meaning it has in relation to its literal meaning. As Dobrovol'skij & Piirainen (2005: 80f) point out, motivation is a subjective criterion that cannot be verified or operationalized in a strict way. One has to take into consideration the fact that motivation depends on one's associations, experience, specific encyclopedic knowledge associated with the literal meaning – something which Langlotz (2006: 51) himself admits, as well as knowledge of an idiom's etymology (although in some cases it can be unrelated to the idiom's current meaning<sup>149</sup>) or ability to relate the literal meaning to a specific context (e.g. the context of boxing for the Finnish idiom *HEITTÄÄ PYYHE KEHÄÄN* lit. 'to throw the towel into the ring', id. 'to give up'), etc. All of the above factors obviously vary from speaker to speaker, thus there is no way to prove how every individual processes a given idiom, nor how his/her ability to motivate it is reflected in his/her use of this idiom. This high degree of subjectivity makes the notion of motivation practically inapplicable for making predictions

---

<sup>149</sup> E.g. Nikanne (personal communication) have pointed out that the meaning of the Finnish idiom *KAKSITERÄINEN MIEKKA* lit. 'two-edged sword', id. 'something which has both positive and negative consequences' could hardly be motivated on the basis of its original Biblical context where it is compared to the word of God: *For the word of God is living and powerful, and sharper than any two-edged sword, piercing even to the division of soul and spirit, and of joints and marrow, and is a discerner of the thoughts and intents of the heart.* (Hebrews 4:12)

about the systematicity of variation, unless one succeeds in finding other sources of motivation which are constant and objective for every speaker.

This is what Langlotz (2006: 66-74) attempts by introducing conceptual patterns of semantic extension as the basis of motivation. These patterns are conceptual metaphors and conceptual metonymy (Lakoff & Johnson 1980; Lakoff 1987), conceptual integration through blending (Fauconnier & Turner 1996, 1998, 2002; Fauconnier 1994, 1997), emblems (a substitute for the corresponding notion of symbol<sup>150</sup>, as in Dobrovol'skij & Piirainen 1997, 2005) and the "cognitive interaction" of all the above. In Langlotz's (2006) semantic analysis, conceptual metaphors are clearly dominant. Langlotz's (2006) idea of idiom motivation is to a very large extent shaped by and dependent on the postulates of the cognitive theory of metaphor developed by Lakoff and his colleagues (especially Lakoff 1987). In Chapter 2 Section 2.3.3, I have presented some essential points of criticism that could be levelled against this theory. Taking them into consideration, there is a good reason to doubt that its models can explain systematic variation.

Another obvious problem with this approach to the notion of systematic variation is that it fails to explain the evidence of usual variation in unmotivated idioms, like *TO KICK THE BUCKET* in Examples (92) – (101) above. Because they are recurrent, examples like these have to be considered as usual and therefore cannot be regarded as wordplay, which, according to the classification presented by Langlotz (2006: 204), is a subcategory of occasional variation and does not extend to the realm of usual variation. Langlotz (2006: 284-285) labels such examples as controversial data and admits that since "there is no obvious conceptual basis in relation to which these patterns [...] could be explained" his theoretical model has no other means to treat them, except for suggesting that:

[...] some *restricted* types of idiom variation with opaque idioms could be subject to the entrenchment of constructional schemas that guide the constrained range within which the alterations can be produced (Langlotz 2006: 285).

---

<sup>150</sup> Lakoff and Johnson (1980: 40) treat cultural and religious symbolism as special cases of metonymy, e.g. *DOVE FOR HOLY SPIRIT*. Dobrovol'skij & Piirainen (2007: 97, 265) also point out that in most cases symbols constitute a metonymical shift.

In other words, this implies that these variants are created and constrained by particular constructions, e.g. *COME A INTENSIFIER CROPPER* for (105) below:

- (105) *DURHAM Squash Club's challenge for the Durham and Cleveland First Division championships has **come a catastrophic cropper*** (ibid. 284).

The explanation seems plausible, but the same constructional schema could in principle apply to motivated idioms as well. In any case, a particular constructional schema does not constrain the range within which the alterations can be produced except for those it represents, so there is no way to determine all possible variations simply by postulating schemas. There have to be more general principles that transcend specific constructions.

#### 4.2.2.3.2 Contextual effects: intentionality and interpretation

The following distinctive criterion that Langlotz (2006) uses when he speaks about systematic variation vs. wordplay is the speaker's context-specific discursive intention to create some "weakly implicated" semantic or stylistic effects, which the hearer would interpret as "strikingly conspicuous". Thus, idiom variation must be regarded as wordplay if it creates such effects and as systematic if it does not (Langlotz 2006: 196f). Langlotz (2006) also refers to relevance theory (Sperber and Wilson 1986) when he compares systematic idiom variation with the notion of strong communication, where:

[...] the speaker primarily engages in the strong communication of the idiomatic meaning without leaving it to the hearer to generate a series of additional weak implicatures (Langlotz 2006: 196).

Non-systematic idiomatic wordplay on the contrary:

[...] creates contextual effects that go well beyond the evocation of the conventional idiomatic meaning: they trigger the weak communication of additional poetic and stylistic effects (Langlotz 2006: 196).

It is not entirely clear how these notions of strong and weak communication which were originally developed for non-figurative language could be applicable to idiom variation. Idioms are figurative expressions with an intrinsic ability to generate poetic effects<sup>151</sup> per se, i.e.

---

<sup>151</sup> Sperber and Wilson (1986: 222) define the term poetic effect as "the peculiar effect of an utterance which achieves most of its relevance through a wide array of weak implicatures".

even if they appear in their conventional form and their conventional meaning is strongly communicated. Apparently, weak implicatures can vary depending on the context, i.e. they are “contextual” in any case.

Obviously, two conditions have to be met in order to produce the effects: firstly, they must really be intended by the speaker and secondly, the hearer must be able to recognize and interpret them according to the speaker’s intention. How should one proceed in distinguishing systematic variation from wordplay in case there is a mismatch between the intention and the interpretation, i.e. certain effects could be intended by the speaker, but since they were weakly implicated, they were not recognized by the hearer? How do we know what effects were intended and if they were intended in the first place? Langlotz (2006: 198) himself admits that recognition and interpretation of weak effects depends on their subjective evaluation by the receiver. Another problem is the interplay of these effects with two other criteria of systematic variation named by Langlotz (2006), namely motivation and recurrence. Should the variant be regarded as wordplay if it creates additional effects but is otherwise recurrent and in complete agreement with the idiom’s motivated semantic structure? Examples (106) and (107) below can be used to illustrate the controversy:

(106) *The piper wants to be paid* (Gazdar et al. 1985: 241).

(107) *To coin a phrase, Lewis’s war-time broadcasts on behalf of God put a pigeon among the cats.* (Langlotz 2006: 197)

Schenk (1995: 257) claims that (106) above is a case of wordplay because it is “intended to be funny”. However, Langlotz (2006: 195) does not see any strikingly humorous effect. According to him, the variant is systematic as long as the criterion of motivation is concerned: “manipulation is in full accord with the motivated semantic structure of the idiom”. On the other hand, Langlotz (2006) claims that example (107) above must be regarded as wordplay because it creates a new, unconventional, context-specific and striking meaning adaptation, while the phrase *TO COIN A PHRASE* explicitly marks its “playful nature”. This may be true as well, but one could still argue that the idiom *TO PUT A CAT AMONG THE PIGEONS* also has a motivated and analyzable structure and that the reversed variant is not as unique and unconventional as it might seem, e.g. searching in Google Groups for the exact string “*pigeon among the cats*” one can find examples like (108) – (113) below.

- (108) *Sorry to make the group seem like alt.cascade, but \*-w - o - w-\* didn't my post put the pigeon among the cats!*<sup>152</sup>
- (109) *Gods only know what would have happened if the original \*hadn't\* been destroyed, only stolen, and Dee had thought of coming up with it: to have two rival Scones would really have set the pigeon among the cats.*<sup>153</sup>
- (110) *To add some equal rights to this discussion and throw a pigeon among the cats, I nominate Grant Watson as a Sexy SciFi Babe. :)*<sup>154</sup>
- (111) *Real news at last! PES6 and FIFA 360 will be next-gen exclusives for 12 months. That'll put the killer pigeon among the cats! "Xbox 360 owns football," intones Lewis.*<sup>155</sup>
- (112) *YES YOU CAN email me! NOW, There's a pigeon among the cats!*<sup>156</sup>
- (113) *Don't worry about being a cat among the pigeons -- you're probably a pigeon among the cats, given the way we tend to get going on this subject!*<sup>157</sup>

Examples (108) – (113) above also demonstrate that the variant itself can be the subject of further variation, e.g. verbal variants (*put/set/throw*), modification (*killer pigeon*) as well as verbless constructions in (112) and (113) above. Data like this challenges the traditional view of variation as application of formal and semantic changes to the idiom's base-form. The problem of the base-form approach will be discussed in more detail in Section 4.4 of this chapter.

To summarize the point of criticism with regard to the criterion of stylistic conspicuousness vs. inconspicuousness: it cannot work as a measure to distinguish between systematic variation and wordplay if the scholar relies solely on his own subjective evaluation and interpretation. In a similar manner, one cannot be sure whether the speaker/writer did or did not have the intention to trigger particular effects. Langlotz (2006: 208) admits that there is no clear-cut threshold between conspicuousness and inconspicuousness and suggests a relative gradability of this criterion. However, in the absence of any explicit parameters that would define the concept of conspicuousness as well as any explanation of how it can be empirically tested, there is practically no way to determine the grade of conspicuousness produced by any particular variant.

<sup>152</sup> GG: rec.arts.mystery/msg/97a07292408f7e65, Sep 20 1995

<sup>153</sup> GG: alt.books.pratchett/msg/0836bac3f6936d12, Jun 21 2000

<sup>154</sup> GG: aus.sf/msg/a16543dbe965a1ea, Sep 7 1998

<sup>155</sup> GG: uk.games.video.xbox/msg/3e1d43b846ee6513, Aug 23 2006

<sup>156</sup> GG: uk.telecom/msg/44ecd6ae2f5b2313, Apr 12 1998

<sup>157</sup> GG: alt.support.menopause/msg/b63739a9ba06bb62, Nov 2 1997

#### 4.2.2.3.3 Communicative motivation and idiom-variation principles

Finally, Langlotz (2006: 12, 205) tries to establish a distinction between systematic idiom variation and idiomatic wordplay relative to several idiom-variation principles, i.e. different strategies that underlie the variation of idiomatic constructions in discourse. These principles are claimed to represent patterns of idiom adaptation in response to communicative purposes. The following five principles are discussed by Langlotz (2006: 205-215):

1. Constructional adaptations
2. Literal-scene manipulation
3. Topic indication
4. Topic-related literal-scene manipulation and conjunction variation
5. Ambiguation and punning

Constructional adaptation is claimed to be an entirely systematic idiom variation principle. However, the criteria which are used to determine which variants are covered by this principle have practically nothing to do with their communicative function. Instead, Langlotz (2006: 206) defines the principle of constructional adaptation in terms of regularity and grammaticality: systematic variants are “usual and conventional grammatical variants that frequently recur in discourse”.

The second principle is literal-scene manipulation. Adaptations of the idiomatic meaning triggered by this principle are contextually motivated (Langlotz 2006: 207). Variants function to elaborate and intensify the idiomatic meaning by describing the literal scene in more detail, e.g. the premodifier *narrow* in (114) below is consistent with the literal meaning, it specifies it and at the same time the idiomatic meaning is intensified:

- (114) *The Chancellor **had a narrow tightrope to walk** and he managed to please a variety of people* (Langlotz 2006: 207).

Literal-scene manipulation is claimed to cover alternations that range from fully systematic variation to wordplay. A new concept of elaborative distance is introduced for the evaluation of a literal-scene manipulation variant as wordplay. The term itself is borrowed from Langacker (1987), but is used by Langlotz (2006) in a different way than the original notion of elaborative distance. In Langacker’s (1987: 68-69) theory it was used to describe a relationship between a superordinate node and a subordinate node in a taxonomic hierarchy. The subordinate

structure (e.g. a hyponym OAK) elaborates the superordinate one (its hypernym TREE), i.e. adds more precision and detail to it. Langacker (1987) calls the degree of added precision elaborative distance. According to Langlotz (2006: 209), elaborative distance is a conspicuous and striking degree of the variant's deviation from the base form. Thus, based on the criteria of conspicuousness and weak implicatures (previously discussed and criticized in Section 4.2.2.3.2), elaborative distance does not add any fundamentally new criteria to the wordplay category. As for the systematicity of literal-scene manipulation, Langlotz (2006) again relates it to the metaphorical motivation of an idiom.

The third principle, called topic indication, in contrast to literal-scene manipulation, is not compatible with the literal meaning, but instead specifies only the idiomatic meaning and relates the idiom to the context. For instance, although the premodifier *financial* in (115) below has to be interpreted literally, it cannot modify the tightrope in its literal sense and thus has to be considered as a figurative-level modification:

- (115) *That sum may seem like a lot of lei (the Romanian currency that purchases next to nothing abroad) but it still left the Romanians **treading a financial tightrope*** (Langlotz 2006: 210).

Langlotz (2006: 211) describes topic indication as a fully systematic type of occasional variation, rather than a form of wordplay, because:

[...] the topic indicator is integrated into the idiom's formal structure according to general grammatical rules for adnominal modification and lexical substitution.

This is in contradiction with his earlier remark that considering the literal meaning of *financial tightrope*, the semantic contribution of the topic indicator *financial* is ungrammatical and cannot be understood as an inherent, qualifying adjective (Langlotz 2006: 210). In this case, modification by *fraying* and substitution by *tottering* in the literal-scene manipulation example (116) below are even more grammatical, since they are not only integrated according to the rules of grammar but also create a consistent literal meaning. Nevertheless, the example is classified by Langlotz (2006: 209) as wordplay by virtue of its striking degree of deviation from the base form:

- (116) *Only the utter ruthlessness of one ravaged, machine-sustained tyrant and the overstretched forces of his fierce yet fragile Imperium kept the human race **tottering along its fraying tightrope*** (Langlotz 2006: 208).



By and large, it remains unclear why Langlotz (2006) cannot regard the fully grammatical literal-scene manipulation variant in (116) above as systematic, while the topic indication variant in (115) above is considered by the author to be fully systematic. Apparently, since both of them are grammatical, but only the latter is regarded as systematic, grammaticality cannot serve as a distinctive feature in this case. As for the “strikingness” of deviation, it is obviously a very subjective criterion.

The variational strategy applied in the fourth variation principle, called topic-related literal-scene manipulation, involves using a linguistic unit that can be associated with both the usage-context and the literal meaning. Unlike literal-scene manipulation, it does not alter the idiomatic meaning and cannot be ascribed an idiomatic interpretation. Unlike topic indication, it is fully compatible with the literal meaning. Like topic indication, this strategy relates the idiom to the context, but in this case it is done via the idiom’s literal meaning (Langlotz 2006: 211-213), e.g. in (117) below:

(117) *Bruce, a shark, found it a part he could really **sink his three rows of teeth into***<sup>158</sup> (Langlotz 2006: 212).

This variation principle is classified as wordplay only, although it is considered as a combination of the partly systematic second principle and the fully systematic third principle, for the reason that it is “highly conspicuous, fully context-dependent and non-predictable” (Langlotz 2006: 213). But the latter two criteria are also claimed to be applicable to the systematic topic indication:

[...] topic indication is a context-specific phenomenon. [...] the occurrence of the topic indicator cannot be predicted outside the context of use [...] (Langlotz 2006: 211).

According to Langlotz (2006: 213), the difference of the topic-related literal-scene manipulation from the topic indication is that the former “does not obey systematic rules”, but he does not explain what kinds of rules in particular. Whatever these rules might be, they are certainly not the rules of grammar, since all the examples that he provides are fully grammatical. Finally, this principle is regarded as wordplay because it involves ambiguity: the conjuncted parts of the variant, e.g. *three rows of teeth* in (117) above, have to be interpreted literally, while the rest of the idiom must be understood idiomatically. But the same ambiguity can

---

<sup>158</sup> From an article on the making of the movie “Jaws”.

actually be found in systematic topic indication as well: in (115) above *financial* has to be interpreted as literally related to finances, while *treading a tightrope* is understood idiomatically.

To summarize this critical overview of Langlotz's (2006) categories of systematic variation and wordplay: both notions lack systematic and objective criteria. Langlotz's (2006: 195) claim that the notion of systematic variability can be qualitatively delineated contradicts with his conclusion (Langlotz 2006: 98) that it is theoretically impossible to define a strict demarcation line between systematic variation and idiomatic wordplay. Langlotz (2006) also suggests that the phenomena of wordplay cannot always be strictly distinguished and separated and therefore are a matter of continuum. According to Langlotz (2006: 205):

[...] potential for interpretative indeterminacy illustrates that the boundaries between different idiom-variation classes are open to interpretation and thus become fluent rather than clear-cut.

Disagreeing on this point, I would like to remark that it is not only a matter of fluent boundaries between classes. Even if one does not assume categories in the classical sense and talks of prototype effects with core and periphery members, one will still need some explicit definition for the core and identification procedures for the periphery (Taylor 2003: 75). In the absence of both, the interpretative indeterminacy will always cause obvious difficulties for both ascribing examples to a variation class and arranging them along the supposed continuum.

### 4.3 Phraseological blends

Some authors (e.g. Omazić 2007, 2008; Omazić and Delibegović 2009) apply Conceptual Blending Theory or Conceptual Integration Theory (Fauconnier & Turner 1996, 1998, 2002) to lexical substitution in PUs as well as to amalgamation of two or more phraseological units in one modification and present them as a result of the mechanisms of conceptual integration. According to the theory (Fauconnier & Turner 1998), people construct conceptual structures called mental spaces as they think and use language. General cognitive processes, called blending, operate on these mental spaces as inputs. Input spaces project on to a separate blended space, which inherits partial structure from each of them, but also produces new emergent structures, as demonstrated in Figure 73 below.

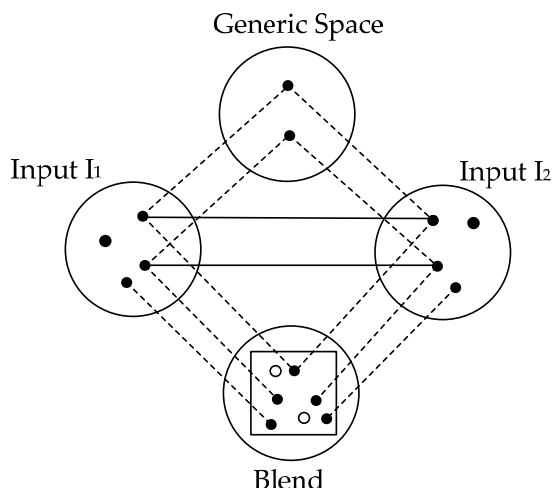


Figure 73 The conceptual integration network (Fauconnier & Turner 2002)

Omazič (2007: 107) argues that a conventional idiomatic expression can provide a frame onto which other input spaces can project other elements of knowledge. The blend inherits the idiom structure and meaning as well as other elements from these input spaces. Omazič (2007) finds the theory very promising from the phraseological perspective, since “many instances of modification to phraseological units may be explained away as instances of blending”.

Conceptual Blending Theory has been criticized for introducing unnecessary complexity into relatively simple linguistic processes (Harder 2003) and because its aspects are not clearly formulated for testing and potential falsification (Gibbs 2000, 2001). To prove that people’s conceptual knowledge somehow predicts the existence of different linguistic behaviour there is a need for empirical, objective evidence (i.e. not based on a theorist’s private intuitions), not just post hoc explanation of people’s linguistic behaviour by postulating theoretical entities such as blending spaces. Ritchie (2004) examines crucial assumptions of Conceptual Blending Theory and shows that its central metaphors (mental spaces, conceptual packets and conceptual blending) obscure processes specified by the model and work against the kind of precise specification that would support meaningful empirical tests.

Although Fauconnier & Turner (2002: 40) relate their model to a network connectionist model of language processing, much of their discussion is inconsistent with the network model. Ritchie (2004) argues

that their metaphors of space, packets and blending work against a network or connectionist understanding of language. The use of circles and boxes in illustrations of the model reinforces the idea of boundaries separating the various conceptual elements and the need for replication of elements within a separate space rather than connection of existing elements in a new composite pattern, as would be entailed by a network model. Richie (2004) concludes that for many of the examples analyzed by Fauconnier & Turner (2002) a simpler and more straightforward analysis seems sufficient where no independent conceptual structures such as generic space and blended space are needed.

In addition to the above-mentioned points of criticism, the Conceptual Integration Theory has another essential drawback – being a very abstract model, it lacks any tools for the analysis of linguistic structure. Omazič (2007: 106), who analyzes the amalgamated form *I fell lock stock and barrel, hook line and sinker, head over heels for you babe*, remarks that it is the syntactic agreement that makes the blend possible: “Without syntactic compatibility between elements the blend would not be operational.” She mentions another blend example *lame duck chase*, which is a combination of the PUs *WILD GOOSE CHASE* and *LAME DUCK*, in which the similarity of the noun phrase structures and a notion of futility expressed by both idioms played a role. But the postulated syntactic compatibility remains unexplained, and so do the alleged semantic link between expressions, since the model itself does not provide us with any apparatus for detailed and explicit analysis of the PU’s syntactic and semantic structure.

#### 4.4 The notion of the base form and approaches to its definition

The notion of the base form of PUs seems to be highly relevant for the study of their variation and variability. As a matter of fact, the very concept of variation cannot be defined without assuming the existence of some kind of initial form to which modifications are applied and to which they can be compared. Cf. the following definitions of some phenomena involved in idiom variation provided by Langlotz (2006: 176-177); each of them is defined via its relation to the base form (boldface is added by me):

Variation/ alteration: any type of formal and semantic change of the **base-form** and/ or idiomatic meaning of an idiom.

Pun variant: an idiom variant that exploits the formal and semantic structure of the **base-form** for the sake of wordplay.

Erroneous variant: a non-intentional departure from the **base-form**; an idiomatic slip-of-the-tongue.

Pseudo variant: a departure from the **base-form** that cannot activate the idiomatic meaning but can only be interpreted literally.

The definition of the base form of PUs is therefore an important methodological issue. First of all, one has to address the question of its status, i.e. in what form and at what level of linguistic representation is it stored? This question can be approached in different ways. According to Chomsky (1980), semantic properties of idioms are exhibited via a base syntactic form at the syntactic level of D-structure, while Bresnan (1982) locates the base form at the level of lexical representation. Both approaches assume that all properties of the idiom are stated once and for all at the level of the single base form, and that all idiom variants can be related to each other via grammatical mechanisms, by which they are derived from the base form. On the other hand, Pulman (1993: 256) asserts that if there are examples of the same idiom appearing in forms that cannot be related by regular syntactic or lexical processes, then there is no syntactic canonical form in which the properties of the idiom can be represented.

The second problem which one is faced with can be formulated as follows: If there is something that can be equated to the base form, which criteria are to be used in order to determine it? In Sections 4.4.1- 4.4.4 I will discuss different approaches to this question.

#### 4.4.1 The lexicographic approach

The concept of the base form is of practical significance to lexicography, i.e. a PU has to be listed in the dictionary according to its base form. In this respect the base form of a PU can be compared to the lemma of a lexeme, which refers to a particular form that is chosen by convention to represent this lexeme in a dictionary, e.g. in many (but not all) languages verbs are conventionally represented by their infinitive form. However, it is not quite clear how this form should be determined for PUs, as they are polylexemic formations that allow other types of variation besides merely morphological inflection of their lexical constituents.

A common practice in idiom variation research is to take dictionary forms of idioms as their base forms. Consider the following quotation by Langlotz (2006: 178):

For *purely practical* reasons, I will equate a given base-form with the idiom's citation-form in idiom dictionaries. I take it for granted that lexicographic practice attempts to record only highly familiar lexicalized constructions belonging to the langue of a given variety (i.e. those units that are entrenched in the mental lexicons of most speakers). Dictionary citation-forms therefore approximate the present view of a usage-based default construction.

A similar statement is made by Naciscione (2001:19): "Practically the base form is the dictionary form [...]." But how practical is it to rely on dictionary forms after all? Apparently, lexicographers are faced with the same problem while choosing the citation form and different dictionaries can, in fact, list different forms of the same PU, e.g.:

- HEITTÄÄ HELMIÄ SIOILLE 'to throw pearls to pigs' (NS; PS; Kari 1993)
- ÄLKÄÄ HEITTÄKÖ HELMIÄNNE SIKOJEN ETEEN 'do not throw your pearls to pigs' (Sinnemäki 1982)
- EI SIKÄ TARTTE HELMIÄ 'a pig does not need pearls' (Laukkanen et al. 1978)
- EI HELMIÄ SIOILLE 'no pearls to pigs' (Parkkinen 2005)

The assumption that a dictionary citation-form approximates the present view of a usage-based default construction has to be empirically tested. For instance, the analysis of empirical data presented in the next section shows that *EI HELMIÄ SIOILLE* 'no pearls to pigs' occurs in 12% of examples and therefore cannot be regarded as a default construction, although it is listed by Parkkinen (2005: 18) as a dictionary form. On the other hand, no contemporary dictionary of Finnish lists *HELMIÄ SIOILLE* 'pearls to pigs' as a citation form, although it is very frequent in the data.

The fact that dictionaries in some cases may not reflect the contemporary language use can also be demonstrated by the following example. Melerovič & Mokienko (2001: 63) in their dictionary, which is based on examples gathered from Russian fictional literature, list the PU *БЫТЬ ПОД БАШМАКОМ* <y> [*byt' pod bashmakóm* <u>] lit. 'to be under <smb's> boot', id. 'to be henpecked, to be under smb's thumb' as the base form and the PU *БЫТЬ ПОД КАБЛУКОМ* <y> [*byt' pod kablukóm* <u>] lit. 'to be under <smb's> heel' as its variant where the lexical constituent *башмак* [*bashmák*] 'boot' is substituted by *каблук* [*kablúk*] 'heel'. Both of these forms are highly familiar to a native Russian speaker, both can be

regarded as lexicalized and entrenched, but the former construction is presented as default and the latter as its variant. This choice could be explained by postulating the higher usage frequency for the PU *БЫТЬ ПОД БАШМАКОМ* <y> [*byt' pod bashmakóm <u>*] lit. 'to be under <smb's> boot' compared to the PU *БЫТЬ ПОД КАБЛУКОМ* <y> [*byt' pod kablukom <u>*] lit. 'to be under <smb's> heel'. However, contemporary language data shows something different.

The search for exact word strings "*под башмаком у*" [*pod bashmakóm u*] (lit. 'under the boot of') and "*под каблуком у*" [*pod kablukóm u*] (lit. 'under the heel of')<sup>159</sup> in Google Groups (3.3.2008) gave 2 and 17 relevant hits respectively. An additional lexico-grammatical search was performed in the Russian National Corpus<sup>160</sup> for the co-occurrence of the preposition *под* 'under' and nouns *каблук* [*kablúk*] 'heel' and *башмак* [*bashmák*] 'boot' in the instrumental case singular with the distance of 0-1 words between the preposition and the nouns. The search gave 22 relevant hits for *БЫТЬ ПОД БАШМАКОМ* <y> [*byt' pod bashmakóm <u>*] lit. 'to be under <smb's> boot' and 24 for *БЫТЬ ПОД КАБЛУКОМ* <y> [*byt' pod kablukom <u>*] lit. 'to be under <smb's> heel'. These numbers could indicate that both forms are nearly equal, if it were not for the difference of their distribution in time: 79% (N=19) of examples with the form *под каблуком* [*pod kablukom*] lit. 'under the heel' were dated from the 1990s and 2000s i.e. can be regarded as representing contemporary language, two examples are from the 1970s, one from 1960s, one from 1920s and only one from the 19<sup>th</sup> century. The situation is different for the form *под башмаком* [*pod bashmakóm*] lit. 'under the boot': 45% (N=10) of examples are from the 19<sup>th</sup> century, another 45% (N=10) are dated from the 1900s to 1940s, one example is from the 1980s and only one from the 1990s. There is not a single example of this form use in the 2000s.

Thus, contemporary language use indicates that *БЫТЬ ПОД КАБЛУКОМ* <y> [*byt' pod kablukom <u>*] lit. 'to be under <smb's> heel' should be regarded as the default form, although the dictionary (Melerovič & Mokienko 2001: 63) presents it as a variant of *БЫТЬ ПОД БАШМАКОМ* <y> [*byt' pod bashmakóm <u>*] lit. 'to be under <smb's> boot'. This example

<sup>159</sup> The preposition *у* 'of, at' was included in the search string in order to maximize precision and reduce the number of hits where the searched phrases could appear in their literal sense. The verb was excluded as the most varying part. All retrieved examples were checked for their relevancy and for duplicate messages.

<sup>160</sup> <http://www.ruscorpora.ru/index.html>

clearly demonstrates that the base form status of a construction can vary diachronically while a dictionary may reflect an out-of-date situation. For further discussion of the diachronic approach to the definition of base form see Section 4.4.4 below.

#### 4.4.2 The quantitative approach

Apparently, the lexicographic approach to the definition of the base form is problematic. But as will be demonstrated below, so is the purely quantitative approach, which would regard the base form as the one that has the highest number of occurrences among other forms in discourse. Thus, Langlotz (2006: 177) defines an idiom's base form as "an idiom's context-independent default-structure that is distilled from various usage-events". Langlotz does not specify what exactly is meant by such "distillation". It may be the case that such definition presupposes that of all the potential forms that an idiom can adopt in discourse, one is "more usual". As I see it, the endeavour to find a more usual form out of all possible less-usual ones presents a major challenge for a corpus-based analysis. I have already discussed this problem in Section 1.3 of the Introduction and pointed out that finding all potential PU variants is a very non-trivial task.

I will now turn to empirical data<sup>161</sup> consisting of 480 tokens of the Finnish PU *HEITTÄÄ HELMIÄ SIOILLE* lit. 'throw pearls to pigs', id. 'cause a transfer of some entity, evaluated by the speaker as good, to a recipient, evaluated as inadequate in some way' and first test the assumption that the base form, and therefore the most usual form, of the PU in question is in fact the construction *X HEITTÄÄ HELMIÄ SIOILLE* | *NP<sub>SUBJ</sub>[N{X}] V{heittää 'throw'} NP<sub>OBJ</sub>[NP<sub>PL</sub>{helmi 'pearl'}] PP<sub>ALL</sub>[NP<sub>PL</sub>{sika 'pig'}]*, as suggested by the dictionary (e.g. Kari 1993). Thus, in order for a PU token to be considered a base form token in a strict sense at least the following major criteria have to be met simultaneously:

- Syntactic construction, e.g. *<NP<sub>SUBJ</sub>>Vtr NP<sub>OBJ</sub> [N<sub>1</sub>] PP<sub>ADJUNCT</sub>[NP[N<sub>2</sub>]]*.
- Constituent morphology, e.g. *NP<sub>PTV\_PL</sub>[N<sub>1</sub>]* and *PP<sub>ALL</sub>[NP<sub>PL</sub>[N<sub>2</sub>]]*.
- Constituent phonology/LCS (e.g. *V{heittää 'throw'}*, *N<sub>1</sub>{helmi 'pearl'}* and *N<sub>2</sub>{sika 'pig'}*).

---

<sup>161</sup> The data gathered from Google Groups is described in Section 1.3 of the Introduction.



Table 4 below presents the supposed default value<sup>162</sup> of each of these formal requirements.

*Table 4 Default values for HELMIÄ SIOILLE base form requirements*

Formal requirement			Default
Syntactic construction			<NP <sub>SUBJ</sub> >Vtr NP <sub>OBJ</sub> [N <sub>1</sub> ]PP <sub>ADJUNCT</sub> [NP[N <sub>2</sub> ]]
Constituent morphology	NP[N <sub>1</sub> ]	CASE	PTV
		NUM	PL
		CASE&NUM	PTV PL
	PP[NP[N <sub>2</sub> ]]	CASE	ALL
		NUM	PL
		CASE&NUM	ALL PL
Constituent phonology/ LCS	V		<i>heittää</i> ‘throw’
	N <sub>1</sub>		<i>helmi</i> ‘pearl’
	N <sub>2</sub>		<i>sika</i> ‘pig’

The actual occurrence numbers are presented in Table 5 below. The two rows marked with solid grey colour contain defaults for the morphosyntactic pattern and the verb phonology/LCS. The average value of these defaults is 35%, while the average of the remaining defaults, which together represent morphology, phonology and LCS of the noun constituents across all constructions in the family, is 88%. For now, I leave open the question of whether these weaker defaults have to be excluded from the picture. I will return to the notion of default once more in Section 5.3.3 of Chapter 5.

*Table 5 Frequencies of occurrence of different criteria for the default form in the HELMIÄ SIOILLE data*

Formal requirement		%	Tokens
<NP <sub>SUBJ</sub> >Vtr NP <sub>OBJ</sub> [N <sub>1</sub> ] PP <sub>ADJUNCT</sub> [NP[N <sub>2</sub> ]]		36	173 out of a total of 480
NP[N <sub>1</sub> ] morphology	PTV	81	385 out of a total of 475 <sup>163</sup>
	PL	92	435 –“–
	PTV PL	79	374 –“–
<PP>[NP[N <sub>2</sub> ]] morphology		92	419 out of a total of 457

<sup>162</sup> The notion of default value has been previously discussed in Section 3.4.2 of Chapter 3.

<sup>163</sup> A different total number for each constituent is acquired by excluding tokens where a construction does not include this constituent.

	PL	92	419 –“–
	ALL PL	85	390 –“–
V{heittää ‘throw’}		34	92 out of a total of 270
N <sub>1</sub> {helmi ‘pearl’}		94	448 out of a total of 475
N <sub>2</sub> {sika ‘pig’}		92	420 out of a total of 457

Each of the defaults in Table 5 above taken separately occurs more frequently in the data than any other non-default realization of the respective category. However, by applying all of the above criteria together to the data one is left with only 48 tokens, i.e. only 10% of the tokens for this PU are tokens of the PU’s base form in the strict sense. If one starts removing some restrictions, one is inevitably faced with the question of how long one is allowed to proceed in order to arrive at a single form, which meets the frequency requirement and the institutionalization requirement. This question cannot be answered within the limits of the quantitative paradigm, and requires a qualitative approach, which will be discussed in Section 4.4.3 below. If one sticks to the requirement that all criteria have to be satisfied simultaneously, one will get another candidate for the position of the default form. In my data, it is the verbless construction *HELMIÄ SIOILLE* | *NP<sub>PTV PL</sub>[N<sub>1</sub>{helmi ‘pearl’} PP<sub>ALL</sub>[NP<sub>PL</sub>[N<sub>2</sub>{sika ‘pig’}]]]* | lit. ‘pearls to pigs’, like in (118) below:

(118) *Sinänsä harmi, 4MB:lla 1280x1024x16bpp riittäisi minulle mukavasti, 24bpp olisi (minusta puhuttaessa) helmiä sioille.*<sup>164</sup>

lit. ‘As such it’s a pity, 280x1024x16bpp with 4MB would do for me nicely, 24bpp would be (talking about me) **pearls to pigs.**’

If compared to the  $\langle NP_{SUBJ} \rangle Vtr NP_{OBJ} [N_1] PP_{ADJUNCT} [NP[N_2]]$ , the morphosyntactic structure  $NP[N_1] PP[NP[N_2]]$  is not a default one: it comes second with 148 tokens (31% out of a total of 480). However, when other criteria (i.e. morphology, phonology and LCS of the constituents) are applied, the remaining number of tokens (N = 125; 26% out of a total of 480) is larger than 48 tokens for the transitive verbal construction  $\langle X \rangle HEITTÄÄ HELMIÄ SIOILLE$  |  $\langle NP_{SUBJ} [N\{X\}] \rangle Vtr\{heittää ‘throw’\} NP_{OBJ PTV PL} [N_1\{helmi ‘pearl’\}] PP_{ALL} [NP_{PL} [N_2\{sika ‘pig’\}]]$ , which we supposed was the default form in the first place. Also there is always a possibility that different data sources will give us different results, e.g. as has already been noted in Section 1.3.1 of the Introduction, 50% of the total of 16

<sup>164</sup> GG: sfnet.atk.linux/msg/f0b07e83b8586673, Mar 26 1996

tokens for the same PU obtained from the Language Bank of Finland (Kielipankki) are represented by the verbless construction.

#### 4.4.3 The qualitative approach

Unlike the quantitative approach, which tries to draw a distinction between the base form and variants as between most frequent forms opposed to less frequent ones, the qualitative approach would characterize the base form as the most optimal one. The “ideal form” has to be reduced to the minimal possible set of constituents in order to avoid redundancy. Cf. Barkema’s (1996: 141) claim:

Lexicalized expressions minimally require the presence of specific lexical items and a specific syntactic structure for their meanings and/or pragmatic functions.

Barkema (1996) defines the base form as “the simplest morphosyntactic form that an expression can take” to activate its specific phraseological meaning and pragmatic function.

It is not clear whether Barkema (1996) speaks of lexicalized expressions (i.e. phraseological units) as out-of-context formations or forms used in discourse. It is also unclear whether he uses the term simplest to denote the shortest possible form. If one assumes that he does and that he actually does not exclude forms used in context from his definition, one is led to the paradoxical discovery that with some contextual support even a single lexical constituent is able to activate the meaning of the whole PU. For instance, it is obvious that the constituent  $NP_{PTV}[N_{PL}\{helmi\} \text{ ‘pearl’}]$  taken outside any context cannot by itself activate the phraseological meaning. However, in Example (119) below it occurs as an isolated phraseological constituent (IPC), while neither the PU nor other lexical constituents are present in the same context. Phraseological interpretation is available even in this case<sup>165</sup>:

- (119) *jos sinä olet uskontovastainen eikä sinua aidosti kiinnosta kuulla mitä buddhalaisuus pitää sisällään, niin voimme lopettaa tämän keskustelun sitten tähän. Tämä on tradition määräämä juttu; ei minun keksintöni. Jeesuksella oli muuten tähän oiva sanonta – siinä oli joitain **helmiä** ja sen semmoisia.*<sup>166</sup>  
 ‘if you are antireligious and not genuinely interested in hearing what Buddhism contains, then we can end this conversation here. This is a thing

<sup>165</sup> A discussion of phraseological constituent isolation and a detailed semantic analysis of the example in (119) can be found in Section 5.4.3 of Chapter 5.

<sup>166</sup> GG: sfnet.keskustelu.uskonto/msg/2c7403e0ea807863, May 1 2002

set by a tradition; not my invention. By the way, Jesus had an excellent saying about this – it had some **pearls** and stuff like that.’

*Helmiä* ‘pearls’ seems to be the simplest (i.e. shortest) morphosyntactic form that *HEITTÄÄ HELMIÄ SIOILLE* ‘throw pearls to pigs’ can take in discourse, but does this entitle one to call it the base form?

Naciscione (2001: 19) points out that the base form is “an abstraction, which has all the most important characteristics of the PU”. This definition of the base form is not free from ambiguity either since it is both inclusive and exclusive at the same time: if one wants to include only the most important characteristics one has to exclude those of minor importance, but on the other hand one would have to make sure that all the most important characteristics are included. It is still not clear which particular characteristics should be regarded most important and which not and what criteria should be used in the evaluation of their importance.

#### 4.4.4 The diachronic approach

Finally, there is a possibility to apply a diachronic perspective and regard the original, historically primary construction as the base form and all later derivatives as its variants. However, there are several problems here as well. Firstly, this option is only available for idioms with identifiable textual sources. Secondly, there are at least two kinds of textual dependence (Dobrovol’skij & Piirainen 2005: 230ff): quotations and allusions. Allusions are references to an entire text or a large passage of text, and in such cases the exact original form cannot be determined. In the case of quotations this could be an easier task. For instance, the Finnish PU *HEITTÄÄ HELMIÄ SIOILLE* lit. ‘throw pearls to pigs’ and the Russian PU *МЕТАТЬ БИСЕР ПЕРЕД СВИНЬЯМИ* [*metát’ biser pèred svín’jami*] lit. ‘cast glass beads before pigs’ both originate from the New Testament Sermon on the Mount, which appears in the Gospel of Matthew chapter 7, verse 6 as a part of a larger utterance ascribed to Jesus:

*Do not give what is holy to the dogs; nor **cast your pearls before swine**, lest they trample them under their feet, and turn and tear you in pieces* (NKJV).

Cf. the following corresponding quotations from several Finnish and Russian Bible translations.

– Old Finnish:

*Älkää antako koirille pyhää, ja älkääät pääarlyjänne heittäkö sikain eteen* [NEG V<sub>IMPV</sub> PL2 *pearl*<sub>PTV</sub> PL 2PL *throw*<sub>IMPV</sub> NEG *pig*<sub>GEN</sub> PL *before*<sub>ILL</sub>], *ettei he niitä joskus tallaa jaloillansa, ja käännä itsiänsä, ja repele teitä* (Biblia, 1776).

– Modern Finnish:

*Älkää antako pyhää koirille, älkääkä heittäkö helmiänne sikojen eteen* [NEG V<sub>IMPV</sub> PL2 -*ka* *throw*<sub>IMPV</sub> NEG *pearl*<sub>PTV</sub> PL 2PL *pig*<sub>GEN</sub> PL *before*<sub>ILL</sub>], *etteivät ne tallaisi niitä jalkoihinsa ja kääntyisi ja repisi teitä* (Vanha suomennos 1933, 1938).

*Älkää antako koirille sitä, mikä on pyhää, älkääkä heittäkö helmiänne sikojen eteen* [NEG V<sub>IMPV</sub> PL2 -*ka* *throw*<sub>IMPV</sub> NEG *pearl*<sub>PTV</sub> PL 2PL *pig*<sub>GEN</sub> PL *before*<sub>ILL</sub>], *etteivät ne tallaa niitä jalkoihinsa ja käy teidän kimppuunne ja raatele teitä* (Uusi suomennos 1992).

– Church Slavonic:

*Не дадите сѣѡмъ псомъ, ни пометајте бисеръ вашихъ предъ свинѣмъ, да не поперѣтъ ихъ ногами своими и вращесѡ расторгнутъ вы.*

*He dadite svätaja psom, ni pometajte biser vaših pred svinijami* [not *throw*<sub>IMPERF</sub> IMPV PL2 *pearl*<sub>GEN</sub> PL *your*<sub>GEN</sub> PL *before* *pig*<sub>INSTR</sub> PL], *da ne poperut ih nogami svoimi i vraššes'a rastórgnut vy* (Elisabeth Bible 1751).

[*Ne dadíte sv'atája psom, ni pometájte biser váših pred svínijami, da ne poperút ih nogámi svoími i vráščšes'a rastórgnut vy.*]

– Modern Russian:

*He davajte svjatyini psam i ne brošajte žemčuga vašego перед свиньями* [not *throw*<sub>IMPERF</sub> IMPV PL2 *pearl*<sub>GEN</sub> SG *your*<sub>GEN</sub> SG *before* *pig*<sub>INSTR</sub> PL], *čtoby oni ne poprali ego nogami svoimi i, obrativšis', ne rasterzali vas* (Synodal translation 1876).

[*Ne davajte sv'atýni psam i ne brošajte žémčuga vášego péréd svínijami, chtoby oni ne poprali ego nogami svoimi i, obrativshis', ne rasterzali vas.*]

As one can see, the set of lexical constituents and their form vary depending on the version. The Finnish PU is closer in its form to the modern translations, while the Russian one has the older translation as its source. The older Finnish translation (1776) contains the archaic word *päärlyjä* 'pearls' instead of *helmiä* 'pearls'. The latter appears both in the 1938 and the 1992 version and also as a constituent in the Finnish PU *HEITTÄÄ HELMIÄ SIOILLE* 'throw pearls to pigs'. The Finnish PU also differs from its source text in the form of the last constituent, which has become conventionalized as the allative phrase *sioille* 'pig<sub>PL</sub> ALL, to pigs' instead of the original postpositional phrase *sikojen eteen* 'pig<sub>PL</sub> GEN before<sub>ILL</sub>, before pigs'.

The Russian PU *МЕТАТЬ БИСЕР ПЕРЕД СВИНЬЯМИ* [*metát' biser péréd svín'jami*] lit. 'cast glass beads before pigs' has inherited its constituent *бисер* [*biser*] 'glass beads', which means 'pearls' in Church Slavonic but has changed its meaning in modern Russian, from the older Church Slavonic version, while the modern translation contains the word *жемчуг* [*žémčug*] 'pearls'. The same is valid for the verbal constituent *метать* [*metát'*] 'cast' in the Russian PU. It is also inherited from the Church Slavonic translation instead of its counterpart *бросать* [*brosát'*] 'throw', which is used in the modern text. This componential structure indicates that the PU in question had become conventionalized in Russian language long before the Synodal translation was made in 1876.

#### 4.4.5 Summary

In Section 4.4 I presented four possible approaches to the notion of the base form of idioms. It turns out that none of them is sufficient to determine which of the forms should be regarded as the base forms. The first definition blindly relies on the dictionary citation form as if it were the ultimate truth. It was demonstrated that even corpus-based dictionaries (Melerovič & Mokienko 2001) can make wrong predictions about the situation in contemporary language. The second definition involves an attempt to determine the most regular pattern out of all possible realizations of an idiom in discourse. This method is clearly dependent on the choice of data source and data sample. There is a risk that one will get several competing forms instead of one base form. The third approach basically lacks a strict definition of the optimal form and any criteria according to which the most important features should be determined. The diachronic approach is only applicable to quotations and therefore cannot be regarded as a general method of base form detection.

### 4.5 The construction family and inheritance network approach to variation

A possible solution to the problem of the base form outlined in Section 4.4 above could be found in the Wittgenstein's (1953) idea of family resemblance. Wittgenstein (1953: 66) points out that there is no need to look for one, essential core meaning of a word, which is common to all uses of that word. Instead one should consider the word's uses as "a

complicated network of similarities, overlapping and criss-crossing". In the same way, instead of looking for one essential base form and its variants one could talk of a network of structurally and semantically overlapping constructions. A similar idea is expressed in the construction family and inheritance approach developed by construction grammarians (e.g. Fillmore 1999, Goldberg 1995, Michaelis 1998, Michaelis & Lambrecht 1996, Penttilä 2006).

In Construction Grammar a grammatical construction is defined as a set of formal conditions on morphosyntax, semantic interpretation, pragmatic function and phonology (Fillmore 1999: 113). A grammar is considered to consist of a structured inventory of grammatical constructions that form a network connected by links of inheritance (Fillmore 1999: 115). Inheritance is a genetic model used to represent formal and semantic correspondences among linguistic expressions (Michaelis 1998). Following Goldberg (1995: 73), if one assumes that two constructions are both syntactically and semantically related, it would mean that they are connected by inheritance links. The relationship of constructions in terms of inheritance is hierarchical: lower level constructions are dominated by higher level ones. An inheritance relation between two constructions  $C_1$  and  $C_2$  such that  $C_2$  inherits from  $C_1$  can be represented as in Figure 74 below.  $C_2$  shares all of the properties of  $C_1$  while adding some of its own. The model also allows for multiple inheritances, i.e. a construction can inherit from more than one dominant construction.

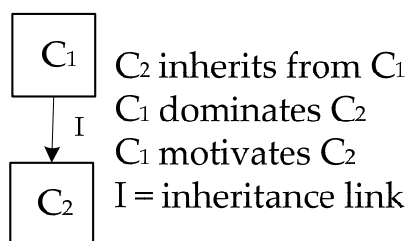


Figure 74 Inheritance (adapted from Goldberg 1995)

The main advantage of the inheritance model is that it promotes generality and economy within the grammar: all nonconflicting syntactic and semantic information shared by two or more constructions need not be listed in the description of each construction. In principle only the dominant construction ( $C_1$ ) needs to be fully specified, while the lower-level construction ( $C_2$ ) can be only partially represented (i.e.

underspecified). The semantic and syntactic specifications associated with construction  $C_2$  will denote only those properties that are not predictable from the properties of construction  $C_2$  (Michaelis 1998: 75).

Inheritance links can describe what kinds of relations exist between the elements of constructions. They are regarded as objects<sup>167</sup> with their own internal structure and are assumed to be related hierarchically (Goldberg 1995: 75). Goldberg (*ibid.* 75-81) distinguishes four major productive types of inheritance links and assumes that there can be also various subtypes<sup>168</sup>:

1. Polysemy links ( $I_P$ ) “capture the nature of the semantic relations between a particular sense of a construction and any extensions from this sense”. Extensions inherit the syntactic specifications of the central sense (*ibid.* 75).
2. Subpart links ( $I_S$ ) are posited when one construction is a proper subpart of another construction and exists independently. Both syntactic and semantic specifications of such a construction are a subpart of the syntactic and semantic specifications of the construction it inherits from (*ibid.* 78).
3. Instance links ( $I_I$ ) are posited when one construction is a special case of another construction, i.e. it is a more fully specified version of the other (*ibid.* 79). An instance link always entails an inverse subpart link (*ibid.* 80-81). Croft (2001: 363) claims that the only type of syntactic relations allowed between constructions are the taxonomic (schema-instance) relations between whole constructions and between an element of one construction and an element of another construction. It means that all relations between constructions are categorizing relations.
4. Metaphorical extension links ( $I_M$ ) are posited between two constructions that are related by a metaphorical mapping (Goldberg 1995: 81).

Michaelis (1998) introduces a pair of inheritance links representing pragmatic contrast:

---

<sup>167</sup> The same idea was earlier expressed by Lakoff (1987).

<sup>168</sup> Lakoff (1987: 420ff) mentions instance, similarity, transformational and metaphorical links that exist between schemas. Goldberg's (1995) notions of instance and metaphorical links are similar to those of Lakoff (1987). Lakoff's (1987) similarity link is posited between two schemas that have shared subschemas (i.e. subparts in Goldberg's (1995) terms).



1. A distality link stands for the semantic extension whereby exponents of past-time reference come to indicate that the speaker is investing a low degree of certainty or confidence in the expressed content or that the speaker views the denoted situation as nonactual.
2. A perspectival-shift link is posited when one expression, A, represents a conventionalized semantic extension of a deictic expression B and that extension consists in the transfer of the deictic reference point to a value not anchored in the speech scene.

Stefanowitsch (2003: 108, 116) also suggests a new type of inheritance link to be added to the apparatus of Construction Grammar: the metonymic link ( $I_{My}$ ), which provides the motivation for the partial structural identity of the indirect speech act constructions (e.g. *Can you X?* used as a request) with the direct construction.

As Leino and Östman (2005: 206) point out, inheritance links are too abstract and therefore can only capture certain rather simple relations between constructions. As their own solution to the problem of variability in grammatical constructions they introduce the notion of metaconstruction – a more abstract or schematic generalization over constructions that captures analogical relationships, systematic similarities and differences of form, as well as a systematic semantic relatedness, which occur between several pairs of constructions and go beyond subsumption and instantiation relations (ibid. 206ff). The authors emphasize that metaconstructions should not be seen as a more abstract, general or schematic level of constructions (ibid. 207), and, in fact, they introduce an alternative way of seeing constructions as resources, where the existence of different levels of schematicity does not enter as a criterion.

Inheritance hierarchy in the sense of Goldberg (1995) and Croft (2001), on the other hand, implies different levels of schematicity and generalization. Constructions are generalizations over actual expressions (constructs). More linguistically instantiated constructions are related via instance links to highly schematic general constructions. General constructions can be compared to phonemes: they are categories and therefore do not receive vocal realization. General constructions dominate one or more specific constructions and express formal and semantic correspondences between those constructions (Michaelis 1998: 129). Michaelis (1998) remarks that the status of general constructions as described by Lakoff (1987) and Goldberg (1995) is somewhat tenuous. She

wonders why such abstract general constructions are not employed in the inheritance network proposed by Lakoff (1987). Instead, Lakoff (1987: 130) represents all overlapping semantic and syntactic specifications of extended constructions as features inherited from the central or immediately dominating construction. In his link diagram Lakoff (1987: 435-436) presents a radial structure where schema 1 occupies a central position; it is the core schema and its instances are of primary importance in the system. Lakoff compares his links to Wittgenstein's (1953) family resemblances, but in fact his network does not conform to the original idea of Wittgenstein's family resemblance model, i.e. that one does not have to presume that there is such a thing as a core sense (or in our case a base form).

Construction networks based on inheritance have been mainly applied to the analysis of general grammatical constructions as well as some idiosyncratic constructions (e.g. *TAKE*  $X_{TEMP}$  and *DO AN*  $N_{PROP}$  in Penttilä 2006). In principle they could also be applied to idioms and idiom variation. As Penttilä (2006: 183) points out, "any aspect of linguistic reality can be taken as the basis of the link between two different constructions". Idiom inheritance links would differ from those of abstract constructions in a way that, in addition to syntactic and semantic features, they would have to capture lexical relations as well.

Variation in lexically partially filled idiosyncratic constructions has been mainly described by assuming an abstract schematic metalevel construction, while considering its different grammatical and lexical variants as more specific instances of the same general phenomenon. The general construction together with its concrete instances forms a construction family (Penttilä 2006: 184). For instance, one can assume that there is in Finnish a general construction  $X OTTI JA Y_{+PAST}$ , lit. 'X took and Y-ed', where  $Y$  is any action performed by  $X$  in the past, while the verb *otti* 'took' imparts this action a tint of unexpectedness or suddenness as well as implies that it was based on a conscious decision (even if the subject is an inanimate object). The relations between this abstract metaconstruction and its concrete instances can be depicted as in Figure 75 below.  $X$  and  $Y$  are the variable elements in the construction. They are marked with subscript indices  $a$  and  $b$ . The same is done to the corresponding parts in the instance constructions.

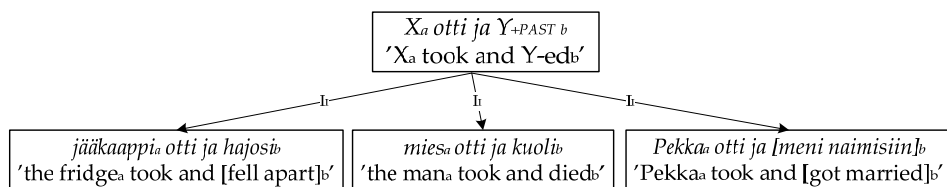


Figure 75 Construction  $X$  OTTI JA  $Y_{+PAST}$ , lit. 'X took and Y-ed' linked to its variants via instance links.

This kind of network can be called taxonomic. It is aimed to describe a relationship of schematicity or generality between different constructions (Croft & Cruse 2004: 262-263). Since instance and subpart links are defined in such a way that an instance link always entails an inverse subpart link (Goldberg 1995: 80-81), any metaconstruction will be simultaneously regarded as a subpart of its variants (Figure 76 below).

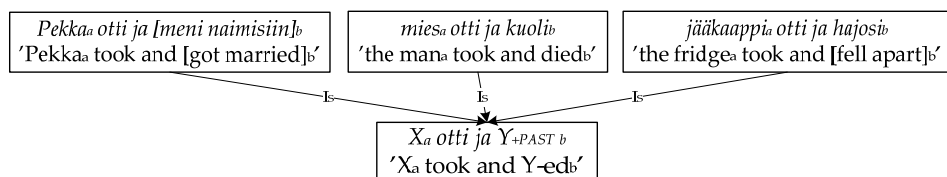


Figure 76 Construction  $X$  OTTI JA  $Y_{+PAST}$ , lit. 'X took and Y-ed' linked to its variants via subpart links.

The assumption that any specific construction is simply an instance of a more schematic construction offers a possibility for construction grammarians to relate different substantive idioms in a taxonomic hierarchy indirectly via higher levels of schematicity. For instance, Figure 77 below represents 4 different Finnish PUs: *HEITTÄÄ HUULTA* lit. 'to throw a lip', id. 'to joke, chatter'; *HEITTÄÄ VEIVINSÄ* lit. 'to throw one's crank', id. 'to die'; *VETÄÄ HIRSIÄ* lit. 'to haul logs', id. 'to sleep' and *VETÄÄ KAULAA* lit. 'to pull a neck', id. 'to break away from competitors (in a race)', which are connected into a taxonomic hierarchy network via the more schematic but verb-specific constructions  $NP_{SUBJ} V\{heittää \text{ 'throw'}\} NP_{OBJ}$  and  $NP_{SUBJ} V\{vetää \text{ 'drag'}\} NP_{OBJ}$ , and on the more abstract level via the wholly schematic construction  $NP_{SUBJ} V NP_{OBJ}$ . A similar taxonomic hierarchy can be found in Croft & Cruse (2004: 264) for the substantive idioms *KICK THE BUCKET* and *KICK THE HABIT*.

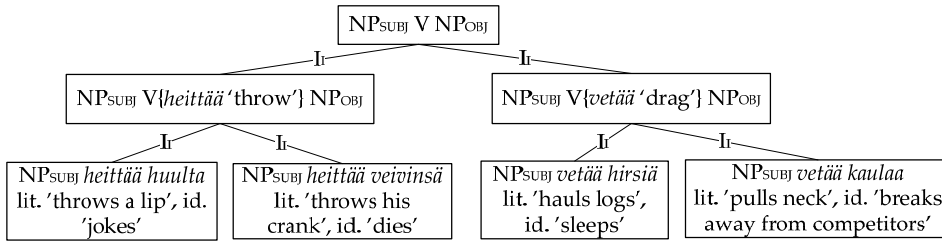


Figure 77 Finnish PUs connected into a taxonomic hierarchy via more schematic constructions

Obviously, this kind of network attempts to relate any idiom with a similar argument structure pattern and verbal constituents, since it only captures syntactic and lexical specifications of the idiom. It can be used for the purposes of syntactic classification, but it is limited in that it does not reflect idiom semantics. Construction grammarians accept the fact that it is typical for a construction to provide only a partial specification of the grammatical structure of its daughter constructions. It is often assumed that even abstract grammatical schemas like  $NP_{SUBJ} V NP_{OBJ}$  specify some semantic information, but to which extent this information is semantic could be of course a matter of discussion. For instance, transitivity is a grammatical category which only counts object arguments of the verbal predicate. When it comes to the idiomatic meaning ('joke, chatter', 'die', 'snore' and 'sprawl'), all of the above idioms have completely different conceptual structures, which are not instances of the verbal constituents' (*heittää* 'throw' and *vetää* 'drag') argument structure, i.e. they have to inherit their meaning from elsewhere. This kind of separate inheritance of structure and meaning is illustrated in Figure 78 below for a group of Finnish PUs with a similar meaning 'to tell jokes, to chatter' and a similar  $NP_{SUBJ} V NP_{OBJ} PTV$  structure: *HEITTÄÄ HUULTA* lit. 'throw a lip', *HEITTÄÄ LÄPPÄÄ* lit. 'throw a flap', *HEITTÄÄ HETULAA* lit. 'throw a whalebone', *HEITTÄÄ LEGENDAA* lit. 'throw a legend'.

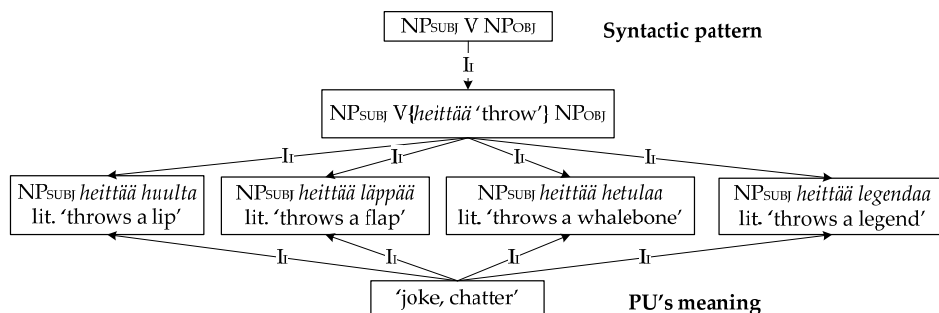


Figure 78 Finnish PUs with a similar meaning 'to tell jokes, to chatter' and a similar  $NP_{SUBJ} V NP_{OBJ}$  PTV structure

In the long run, it is not clear whether we are dealing with inheritance or just a simple ad hoc syntactic and semantic categorization of idioms. Moreover, although general syntactic rules of language (e.g. patterns like  $NP_{SUBJ} V NP_{OBJ}$ ) are analyzed by construction grammarians (Fillmore, Kay & O'Connor 1988) as constructions, they do not have a status of constructions within the framework of Conceptual Semantics, since they correspond to regular linking principles, while Conceptual Semantics treats constructions as linking devices (correspondence rules between different levels of representation) that license irregular linking patterns (Nikanne 2005a: 199).

So far I have shown how different schematic and substantive idioms can be related to each other via taxonomic networks. The problem becomes more complicated when it comes to creating a network of a substantive idiom's variants. In the case of *HEITTÄÄ HELMIÄ SIOILLE* lit. 'throw pearls to pigs', the construction family is not formed by a schematic idiom and its substantive instances like *X OTTI JA Y+PAST*, lit. 'X took and Y-ed' in Figure 75 and Figure 76 above. Instead, we have a group of substantive constructions that are both related and different in some aspect. This kind of constructional variation is not something exceptional: the same phenomenon is typical for many other idioms as well. Croft & Cruse (2004) formulate the reason why these variants should be represented as separate constructions in the following way:

Any construction with unique idiosyncratic morphological, syntactic, lexical, semantic, pragmatic or discourse-functional properties must be represented as an independent node in the constructional network in order to capture a speaker's knowledge of their language. That is, any quirk of a construction is sufficient to represent that construction as an independent node (Croft & Cruse 2004: 263).

In Figure 79 below following constructions are represented in a taxonomic network which is made according to Goldberg's (1995) ideas of inheritance but also using some formalism of Conceptual Semantics:

- *X HEITTÄÄ/ANTAA HELMIÄ SIOILLE* |  $NP_{SUBJ}[N\{X\}]$   $Vtr\{heittää\}$  'throw'/*antaa* 'give'  $NP_{OBJ}[N_{PL}\{helmi\}$  'pearl']]  $PP_{ALL}[NP_{PL}\{sika\}$  'pig']] | lit. 'X throws/gives pearls to pigs'
- *HELMET/HELMIÄ MENEVÄT/MENEE SIOILLE* |  $NP_{SUBJ}[N_{PL}\{helmi\}$  'pearl']]  $Vintr\{mennä\}$  'go'  $PP_{ALL}[NP_{PL}\{sika\}$  'pig']] | lit. 'pearls go to pigs'
- *X POMII HELMIÄ SIKOJEN JOUKOSTA*  $NP_{SUBJ}[N\{X\}]$   $Vtr\{poimia\}$  'pick'  $NP_{OBJ}[N_{PL}\{helmi\}$  'pearl']]  $PP_{GEN}\ JOUKOSTA[NP_{PL}\{sika\}$  'pig']] | lit. 'X picks pearls among pigs'
- *X VARASTAA HELMIÄ SIOILTA* |  $NP_{SUBJ}[N\{X\}]$   $Vtr\{varastaa\}$  'steal'  $NP_{OBJ}[N_{PL}\{helmi\}$  'pearl']]  $PP_{ABL}[NP_{PL}\{sika\}$  'pig']] | lit. 'X steals pearls from pigs'
- *SIAT HEITTÄVÄT HELMIÄ LANTALÄTÄKKÖÖN* |  $NP_{SUBJ}[N_{PL}\{sika\}$  'pig']]  $Vtr\{heittää\}$  'throw'  $NP_{OBJ}[N_{PL}\{helmi\}$  'pearl']]  $PP_{ILL}[NP_{PL}\{lantälätäkö\}$  'dung puddle']] | lit. 'pigs throw pearls into a dung puddle'

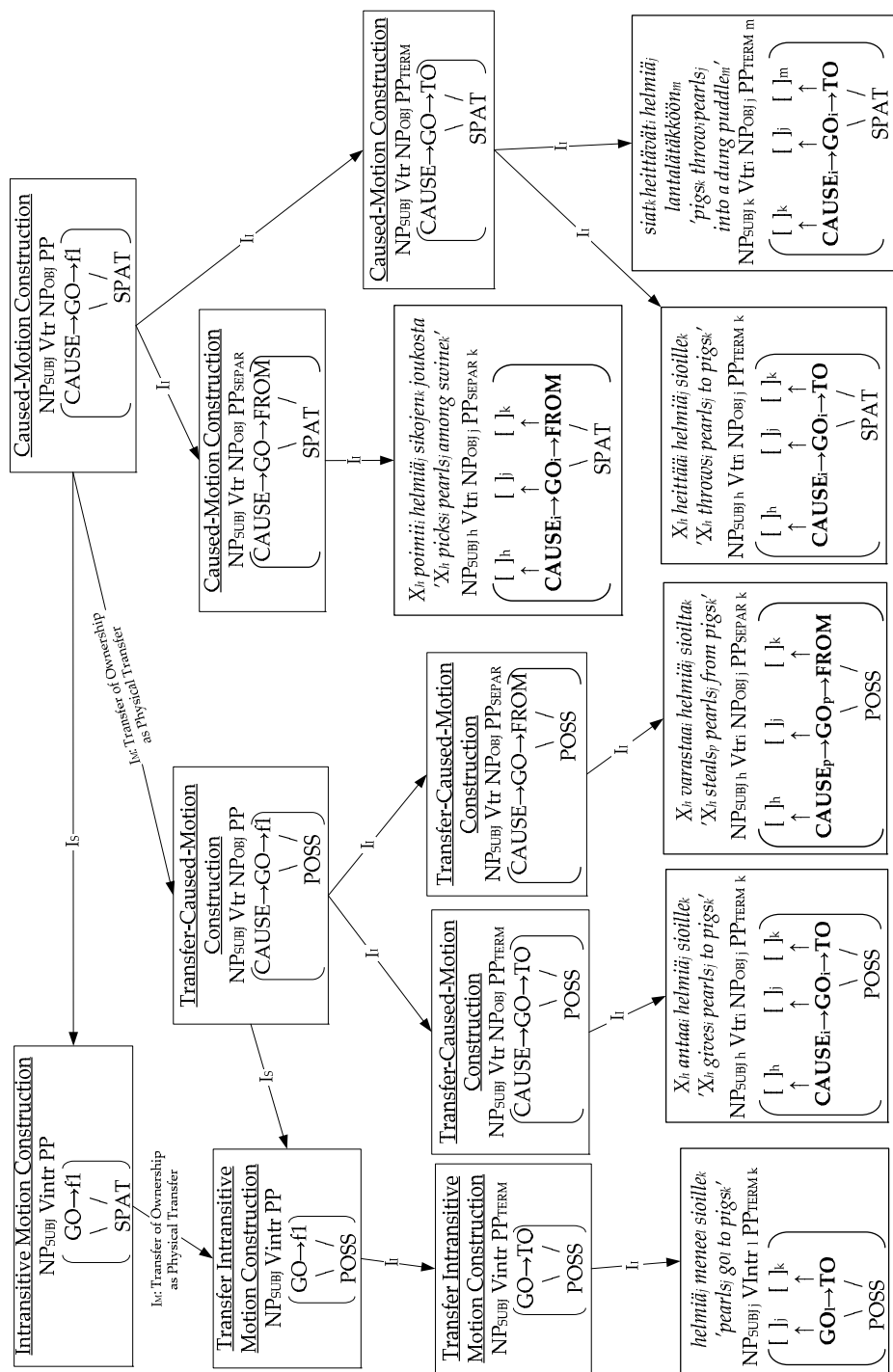


Figure 79 Five HELMIÄ SIOILLE constructions arranged into a taxonomic network.

For the sake of simplicity, PU constructions in Figure 79 are presented in their regular, rule-based meaning (i.e. CS). Adding the context-dependent CS/PU, which involves a high degree of variation, would considerably complicate the picture. This model does not allow linking the idiom variants directly to the base form (whatever it might be). Instead generalizations across these constructions have to be captured by stating them at higher, more abstract nodes in the inheritance hierarchy (Goldberg 1995: 108). The parent node for all constructions is the spatial Caused-Motion Construction ( $NP_{SUBJ} V NP_{OBJ} PP$ ) with unspecified PATH relations (CAUSE→GO→f1). Its two instances are specifications over PATH directions (TO and FROM), while substantive constructions *SIAT HEITTÄVÄT HELMIÄ LANTALÄTÄKKÖÖN* |  $NP_{SUBJ}[N_{PL}\{sika\} \text{ 'pig'}] Vtr\{heittää\} \text{ 'throw'} NP_{OBJ}[N_{PL}\{helmi\} \text{ 'pearl'}] PP_{ILL}[NP_{PL}\{lantälätäkkö\} \text{ 'dung puddle'}]$  | lit. 'pigs throw pearls into a dung puddle', *X HEITTÄÄ/ANTAA HELMIÄ SIOILLE* |  $NP_{SUBJ}[N\{X\}] Vtr\{heittää\} \text{ 'throw'}/\text{'give'} NP_{OBJ}[N_{PL}\{helmi\} \text{ 'pearl'}] PP_{ALL}[NP_{PL}\{sika\} \text{ 'pig'}]$  | lit. 'X throws/gives pearls to pigs' and *X POMII HELMIÄ SIKOJEN JOUKOSTA*  $NP_{SUBJ}[N\{X\}] Vtr\{poimia\} \text{ 'pick'} NP_{OBJ}[N_{PL}\{helmi\} \text{ 'pearl'}] PP_{GEN\ JOUKOSTA}[NP_{PL}\{sika\} \text{ 'pig'}]$  | lit. 'X picks pearls among pigs' are respective instances of these.

The spatial interpretation constitutes an instance of the TO-directional Caused-Motion Construction. However, if the transfer of pearls to pigs is interpreted as being possessive instead of spatial (as in *X ANTAA HELMIÄ SIOILLE* |  $NP_{SUBJ}[N\{X\}] Vtr\{antaa\} \text{ 'give'} NP_{OBJ}[N_{PL}\{helmi\} \text{ 'pearl'}] PP_{ALL}[NP_{PL}\{sika\} \text{ 'pig'}]$  | lit. 'X gives pearls to pigs'), it cannot be directly derived from the Caused-Motion Construction. Goldberg labels such a possessive construction a Transfer-Caused-Motion Construction and argues that it represents a metaphorical extension of the Caused-Motion Construction. The metaphor called Transfer of Ownership as Physical Transfer involves understanding possession as the possessed being located next to the possessor, transferring an entity to a recipient as causing the entity to move to that recipient, and transferring ownership away from a possessor as taking that entity away from the possessor (Goldberg 1995: 89). Just like the spatial Caused-Motion Construction, the general Transfer-Caused-Motion Construction has to be unspecified with regard to its PATH, i.e. whether the entity is transferred to a recipient or taken away from the possessor. Specifications are given in lower level constructions, which in turn instantiate substantive constructions *X ANTAA HELMIÄ SIOILLE* |  $NP_{SUBJ}[N\{X\}] Vtr\{antaa\} \text{ 'give'} NP_{OBJ}[N_{PL}\{helmi\}$



'pearl'})  $PP_{ALL}[NP_{PL}\{sika \text{ 'pig'}\}]$  | lit. 'X gives pearls to pigs' and  $X VARASTAA HELMIÄ SIOILTA$  |  $NP_{SUBJ}[N\{X\}] Vtr\{varastaa \text{ 'steal'}\} NP_{OBJ}[NP_{PL}\{helmi \text{ 'pearl'}\}] PP_{ABL}[NP_{PL}\{sika \text{ 'pig'}\}]$  | lit. 'X steals pearls from pigs'. In addition,  $HELMIA MENE SIOILLE$  |  $NP_{SUBJ} PTV[NP_{PL}\{helmi \text{ 'pearl'}\}] Vitr\{mennä \text{ 'go'}\} PP_{ALL}[NP_{PL}\{sika \text{ 'pig'}\}]$  | lit. 'pearls go to pigs' is also a construction that denotes possessive transfer. Its parent is the Transfer Intransitive Motion Construction, which is a metaphorical extension of the spatial Intransitive Motion Construction, which in turn is a subpart of the Caused-Motion Construction. At the same time, the Transfer Intransitive Motion Construction can be regarded as a subpart of the Transfer-Caused-Motion Construction, which is a metaphorical extension of the spatial Caused-Motion Construction. This overlap could be explained away by multiple inheritance, which allows a given construction to inherit conflicting information from more than one dominant construction (Goldberg 1995: 73, 97), but in the case of the Transfer Intransitive Motion Construction information is the same: it is inherited twice via two different paths but in the reverse sequence.

So far, I have attempted to relate five variants by capturing generalizations of their pared-down syntactic and conceptual structures. The spatial Caused-Motion Construction was a sufficiently high node for that task. However, in order to include more constructions into the network I need to introduce more levels and more detailed representations. Figure 80 below demonstrates that construction  $SIAT SAAVAT HELMIÄ$  |  $NP_{SUBJ}[NP_{PL}\{sika \text{ 'pig'}\}] Vtr\{saada \text{ 'get'}\} NP_{OBJ}[NP_{PL}\{helmi \text{ 'pearl'}\}]$  | lit. 'pigs get pearls' can only be incorporated into the hierarchy via the abstract Transitive Construction node.

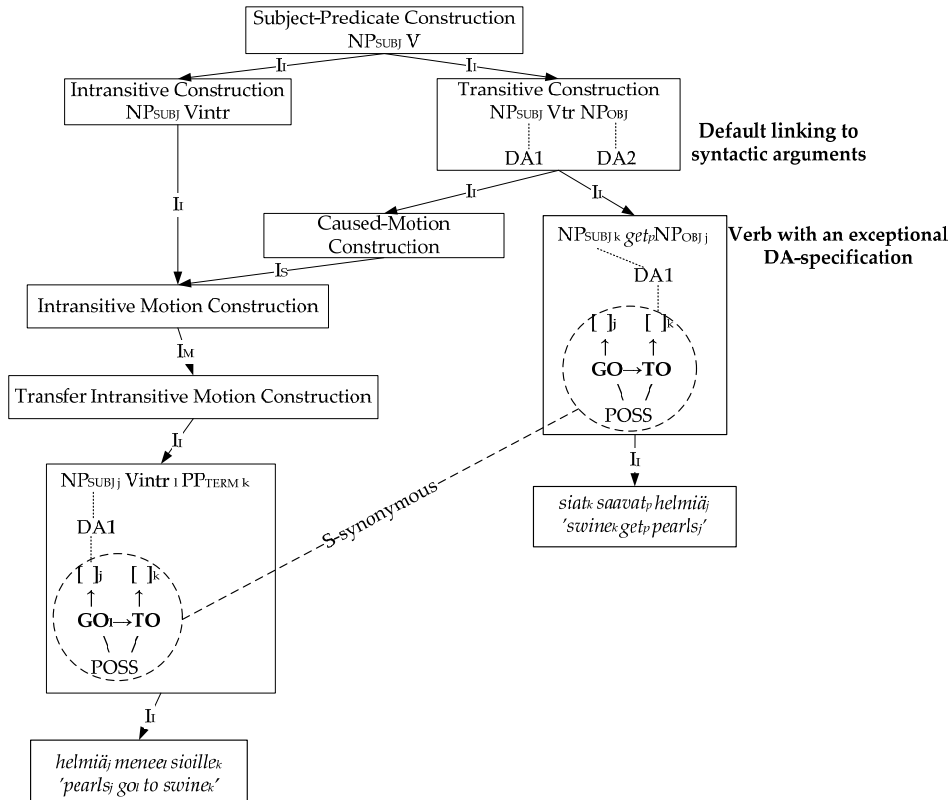


Figure 80 Construction *SIAT SAAVAT HELMIÄ* |  $NP_{SUBJ}[N_{PL}\{sika\} \text{ 'pig'}]$   $V_{tr}\{saada\}$  'get' |  $NP_{OBJ}[N_{PL}\{helmi\} \text{ 'pearl'}]$  | lit. 'pigs get pearls' incorporated into the hierarchy via the abstract Transitive Construction node

Figure 80 above also shows that the lexical argument linking has to be taken into account. The default linking between Direct Arguments (DAs) and syntactic arguments in regular transitive constructions is that DA1 corresponds to subject and DA2 to object. The verb *saada* 'get' violates this linking principle and specifies an exceptional lexical correspondence between the argument level and DA level (Nikanne 2005a: 205; 1997a: 89, 107-108). Cases like this, i.e. when information from the dominated node conflicts with that of a dominant node, do not fit into the complete mode inheritance model (e.g. Kay 1984, Fillmore & Kay 1993), which presumes that all information is inherited. For this reason Goldberg (1995: 73-74) exploits the normal (or default) mode of inheritance, which allows exceptions in lower nodes to override inherited information.

Apart from the difference in argument linking, the transfer intransitive motion construction, where the entity is transferred to a recipient (*HELMIÄ*

MENEE SIOILLE | NP<sub>SUBJ</sub> PTV[N<sub>PL</sub>{helmi ‘pearl’}] Vitr{mennä ‘go’} PP<sub>ALL</sub>[NP<sub>PL</sub>{sika ‘pig’}] | lit. ‘pearls go to pigs’), has a similar conceptual structure to the verb *saada* ‘get’ (SIAT SAAVAT HELMIÄ | NP<sub>SUBJ</sub>[N<sub>PL</sub>{sika ‘pig’}] Vtr{saada ‘get’} NP<sub>OBJ</sub>[N<sub>PL</sub>{helmi ‘pearl’}]) | lit. ‘pigs get pearls’). Since these constructions are not related syntactically, their semantic synonymy does not constitute a motivation link (Goldberg 1995: 91). This is a consequence of the Principle of Maximized Motivation, according to which a construction is motivated by and related to another construction semantically if it is related to that construction syntactically, as well as the Principle of No Synonymy, which states that syntactically distinct constructions must be semantically or pragmatically distinct (Goldberg 1995: 67, quoting Lakoff 1987; Bollinger 1968; Haiman 1985; Clark 1987; MacWhinney 1989).

The verbless construction HELMIÄ SIOILLE | NP<sub>PTV</sub> PL[N{helmi ‘pearl’}] PP<sub>ALL</sub>[NP<sub>PL</sub>[N{sika ‘pig’}]] | lit. ‘pearls to pigs’ forms its own network, since it appears in several different clausal constructions: mostly as a nominal predicative (123), but also in subjectless clauses (122), as an independent disconnected utterance (120), in an object complement *That* clause (121), in comparative *kuin* ‘like, as if’ predicative structures (124) and as a part of a compound (125):

- (120) *Eipä näytä sinulla olevan hitustakaan huumorintajua. Sen verran kurttuotsaisella tosikkomaisuudella vastaat. **Helmiä sioille!***<sup>169</sup>  
lit. ‘It seems that you don’t have even the slightest sense of humour. You answer with such a frowning seriousness. **Pearls to pigs!**’
- (121) *Tuossa on jotain niin syvällisfilosofista viisautta, että alkaa epäillä, että **helmiä sioille**, kun täällä noita lauot.*<sup>170</sup>  
lit. ‘There is some wisdom there which is so profoundly philosophical that one begins to suspect that **pearls to pigs**, when you shoot out those here.’
- (122) *Tuolla Kimmon koneella sais vaikka mitä aikaan, mutta **on** vaan ns. **helmiä sioille**.*<sup>171</sup>  
lit. ‘One could accomplish anything at all with that Kimmo’s engine, but is just so called **pearls to pigs**.’
- (123) *Laitetaan tähän maistiainen vitsiryhmästä, vaikka **se onkin helmiä sioille**.*<sup>172</sup>

<sup>169</sup> GG: sfnet.keskustelu.uskonto.kristinusko/msg/54040c24d8c775f8, May 9 2006

<sup>170</sup> GG: sfnet.huuhaa/msg/10da7fbd2442847d, Apr 16 2006

<sup>171</sup> GG: sfnet.harrastus.mp/msg/1c3c11738bdd3e75, Dec 12 2005

<sup>172</sup> GG: sfnet.huuhaa/msg/dee7f718676010a4, Feb 20 2005

lit. 'Let's put here a sample from the joke group, although it is **pearls to pigs**.'

- (124) *Uskon kuitenkin, että paljon tärkeämpää on kiinnittää huomio sisällön viestinnällisyyteen kuin mennä tekemään sellaisia dokumentteja, jotka nykyisille selaimille ovat kuin helmiä sioille...*<sup>173</sup>

lit. 'I believe anyway that it is more important to pay attention to the communicativeness of the content than to go on making such documents that are like **pearls to pigs** to present-day browsers.'

- (125) *Yleisön mielestä kyseessä on kuitenkin helmiä sioille ratkaisu ja he pitävät itseään niin tärkeänä että vähintään pääkäyttäjää pitää olla.*<sup>174</sup>

lit. 'In the eyes of the public it is nevertheless a question of a **pearls to pigs solution** and they consider themselves so important that they should be at least a main user.'

Figure 81 below demonstrates the network formed by different instances of the verbless construction *HELMIÄ SIOILLE* | *NP<sub>PTV</sub> PL[N{helmi 'pearl'} PP<sub>ALL</sub>[NP<sub>PL</sub>[N{sika 'pig'}]]]* | lit. 'pearls to pigs'.

<sup>173</sup> GG: sfnet.viestinta.www/msg/d054c4237ac19d8c, Nov 22 2005

<sup>174</sup> GG: sfnet.atk.turvallisuus/msg/f5bf63bef569d825, Nov 11 2005



Furthermore, one has to be able to account for substantive constructions where lexical constituents are substituted, inflected, or the number of constituents is reduced or increased. Apparently, these forms will have their own sub-networks. In the inheritance network approach constructions with a reduced number of constituents would be regarded as subparts of larger forms, e.g. Figure 82 below presents  $X$  HETTÄÄ HELMIÄ |  $NP_{SUBj}[N\{X\}]$   $Vtr\{heittää\}$  'throw'  $NP_{OBJ}[N_{PL}\{helmi\}$  'pearl']] | lit. 'X throws pearls' as a subpart of  $X$  HEITTÄÄ HELMIÄ SIOILLE |  $NP_{SUBj}[N\{X\}]$   $Vtr\{heittää\}$  'throw'  $NP_{OBJ}[N_{PL}\{helmi\}$  'pearl']]  $PP_{ALL}[NP_{PL}\{sika\}$  'pig']] | lit. 'X throws pearls to pigs'.

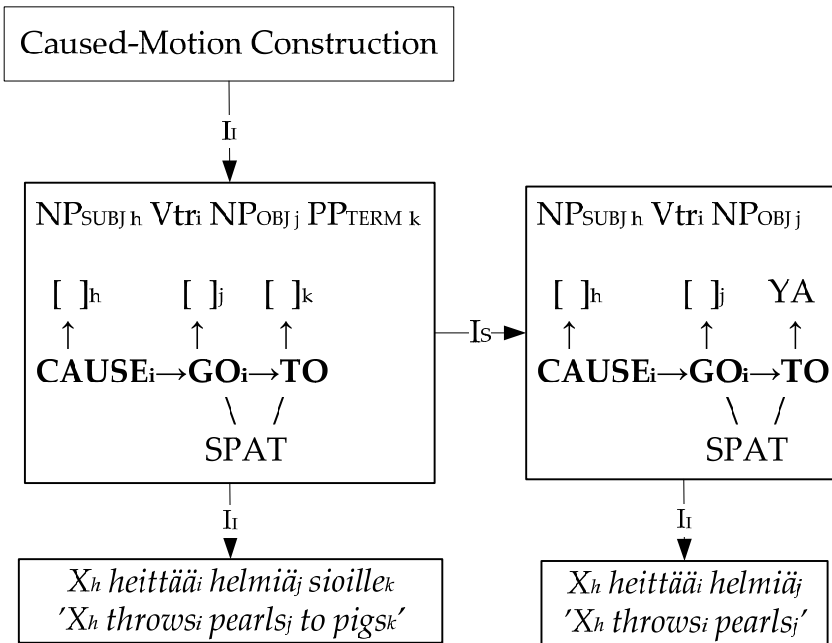


Figure 82  $X$  HETTÄÄ HELMIÄ |  $NP_{SUBj}[N\{X\}]$   $Vtr\{heittää\}$  'throw'  $NP_{OBJ}[N_{PL}\{helmi\}$  'pearl']] | lit. 'X throws pearls' as a subpart of  $X$  HEITTÄÄ HELMIÄ SIOILLE |  $NP_{SUBj}[N\{X\}]$   $Vtr\{heittää\}$  'throw'  $NP_{OBJ}[N_{PL}\{helmi\}$  'pearl']]  $PP_{ALL}[NP_{PL}\{sika\}$  'pig']] | lit. 'X throws pearls to pigs'

The verbless construction  $HELMIÄ$   $SIOILLE$  |  $NP_{PTV\ PL}[N\{helmi\}$  'pearl']]  $PP_{ALL}[NP_{PL}[N\{sika\}$  'pig']]] | lit. 'pearls to pigs' also seems to be partial in the sense that it lacks a verb. However, the analysis of this NP shows that although its lexical and morphological form (*helmiä* 'pearl<sub>PL PTV</sub>', *sioille* 'pig<sub>PL ALL</sub>') are identical to those of the verbal constructions  $X$  HEITTÄÄ  $HELMIÄ$   $SIOILLE$  |  $NP_{SUBj}[N\{X\}]$   $Vtr\{heittää\}$  'throw'  $NP_{OBJ}[N_{PL}\{helmi\}$  'pearl']]

$PP_{ALL}[NP_{PL}\{sika \text{ 'pig'}\}]$  | lit. 'X throws pearls to pigs' and  $HELMIÄ \text{ MENE } SIOILLE$  |  $NP_{SUBJ} \text{ PTV}[NP_{PL}\{helmi \text{ 'pearl'}\}] \text{ V}_{intr}\{mennä \text{ 'go'}\} PP_{ALL}[NP_{PL}\{sika \text{ 'pig'}\}]$  | lit. 'pearls go to pigs', it cannot be analyzed as their subpart due to the differences in syntactic and semantic properties. In verbal constructions, the syntactic functions of  $NP_{PTV}[NP_{PL}\{helmi \text{ 'pearl'}\}]$  are Object (transitive construction) and Subject (intransitive construction) and its thematic role is Theme. The syntactic function of  $PP_{ALL}[NP_{PL}\{sika \text{ 'pig'}\}]$  in both verbal constructions is Adverbial and its thematic role is Landmark (Goal). In the verbless construction,  $PP_{ALL}[NP_{PL}\{sika \text{ 'pig'}\}]$  has to be analyzed as the postmodifier (local case attribute) of the head  $NP_{PL}\{helmi \text{ 'pearl'}\}$  and the NP as a whole constitutes a single thematic argument with the role of Landmark (Location). In the light of these differences in the internal structure, the only plausible way to link  $NP_{PTV}[NP_{PL}\{helmi \text{ 'pearl'}\}] PP_{ALL}[NP_{PL}\{sika \text{ 'pig'}\}]$  lit. 'pearls to pigs' to verbal constructions in the PU's construction family would be via the lexical constituents, morphological form and idiomatic meaning.

According to Goldberg (1995: 98), constituents of constructions are also treated as objects (constructions). They can inherit from and be inherited by other constructions. Lexical constituents *helmi* 'pearl' and *sika* 'pig' are especially relevant for my analysis, since they are shared by most of the constructions in this PU's family and therefore are determinant as criteria of family membership. Figure 83 below presents a simplified (in the sense that no semantic, morphological or syntactic information is taken into account) inheritance network, where constructional nodes are related via subpart links to the lexical constituents. Bold lines indicate that these constituents are inherited from the original Biblical form  $ÄLKÄÄ \text{ HEITTÄKÖ } HELMIÄNNE \text{ SIKOJEN } ETEEN$  |  $NEGV_{IMPV} \text{ PL2 } -ka \text{ V}_{IMPV} \text{ NEG}\{heittää \text{ 'throw'}\} NP_{PTV}[NP_{PL} \text{ 2PL}\{helmi \text{ 'pearl'}\}] PP_{ILL}[NP_{GEN}[NP_{PL}\{sika \text{ 'pig'}\}]] P\{eteen \text{ 'before'}\}$  | lit. 'do not throw your pearls before pigs'.

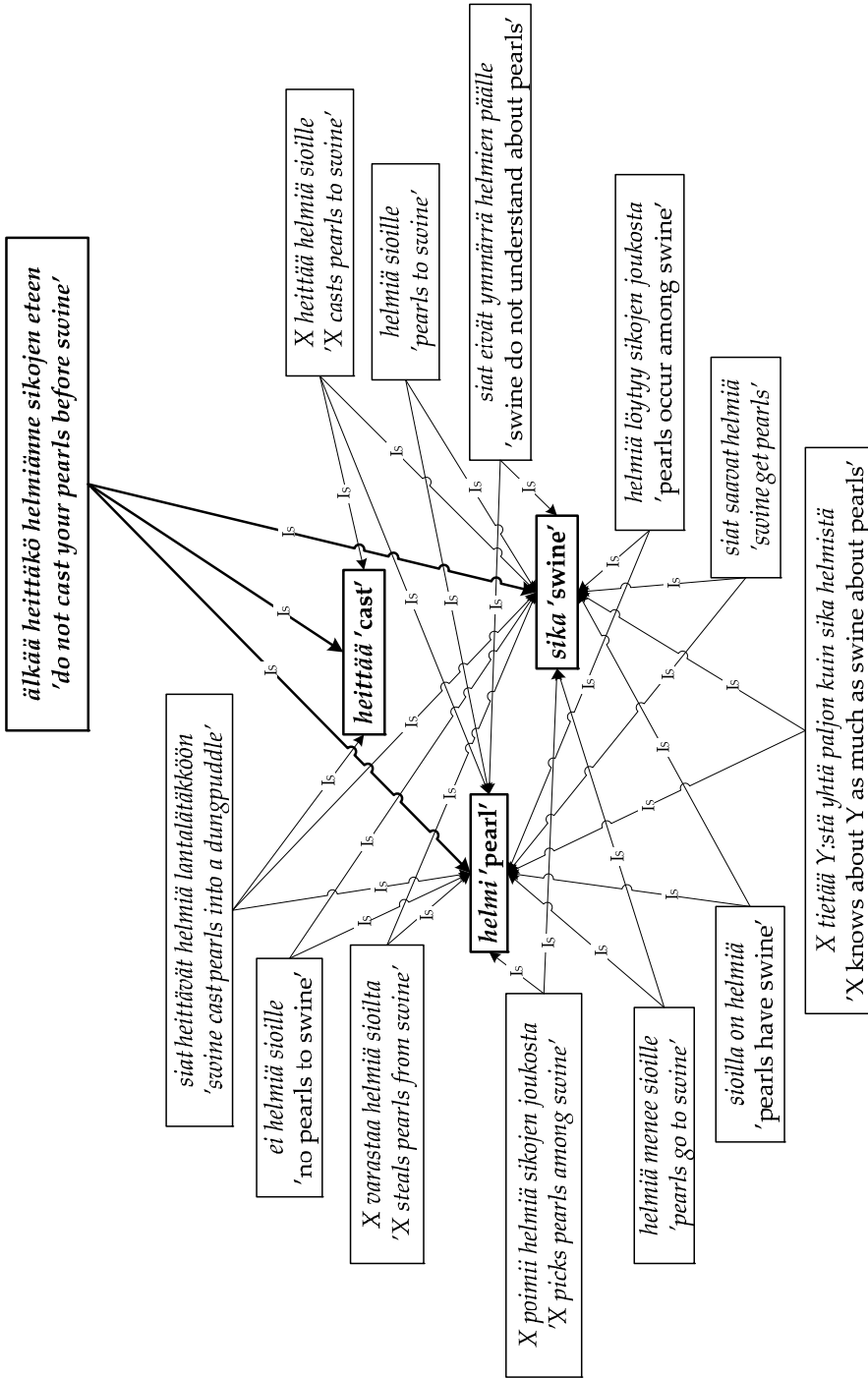


Figure 83 Constructional nodes related via subpart links to the lexical constituents



Lexical constituents can vary in some of the above constructions. I have already mentioned that construction grammars are able to capture lexical variation only by postulating a metaconstruction where the varying part is in some way underspecified, e.g. the word is represented only by its syntactic category and/or function and no phonological form is provided. This method works well for partly schematic constructions like *X OTTI JA Y<sub>+PAST</sub>*, lit. ‘X took and Y-ed’, where at least some part of the idiom is constant and some are variable and these variants are represented as instances of the metaconstruction (see Figure 75 and Figure 76 above). In the case of the *HELMIÄ SIOILLE* construction family, we are dealing with a substantive idiom, where all lexical slots are initially specified. Since all of them are substitutable by other lexical items one will need to underspecify the whole set of constituents in order to get a sufficient generalization, as is done in Figure 84 below. As a result, *HEITTÄÄ HELMIÄ SIOILLE* ‘throw pearls to pigs’ gains the status of just one of the possible instances of the abstract schema *NP<sub>SUBJ</sub> Vtr NP<sub>OBJ</sub> PP<sub>ALL</sub>*. I.e. if one wants to give it a default status, one needs a means to represent default linking. This kind of linking will be presented in the next chapter as a part of the multi-tiered network model.

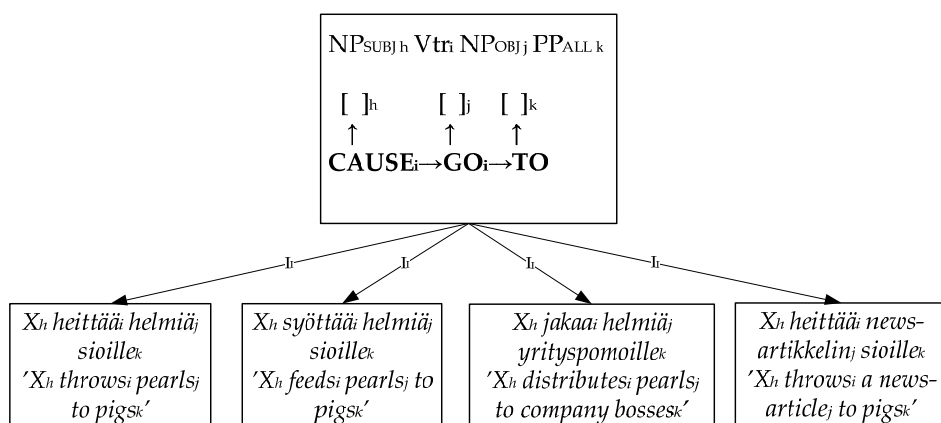


Figure 84 Generalization of lexical instances

Figure 84 above indeed captures the generalization, but the problem is that there is nothing that prevents one from drawing instance links from *NP<sub>SUBJ</sub> Vtr NP<sub>OBJ</sub> PP<sub>ALL</sub>* to expressions like *X antaa lahjan äidille* ‘X gives a present to mother’ or *X heittää pallon pojalle* ‘X throws a ball to a boy’, which perfectly fit the argument structure outlined above, but do not necessarily belong to this idiom family. I.e. apart from generalizations,

one needs a way to stipulate for restrictions in order to exclude irrelevant instances. Croft (2001: 363) claims that prototypes and implicational hierarchies are systematic patterns of variation across constructions that characterize cross-constructional diversity and constrain the distribution and even the form of constructions used for particular functions. However, lexical variation within a PU does not seem to fit into these patterns. A prototype is a privileged subset of members of a category that represent the best exemplars of the category (Croft 2001: 73). Both *heittää* ‘cast’, *jakaa* ‘distribute’ and *syöttää* ‘feed’ in Figure 84 above belong to the syntactic category V, but none of them can be regarded as more prototypical or more peripheral member of this category. *Helmi* ‘pearl’ or *sika* ‘pig’ are not better representatives of a category noun than *artikkeli* ‘article’ and *yrittäjä* ‘company boss’. If one tries to apply conceptual categorization instead of syntactic, one will not even be able to place *helmi* ‘pearl’ and *artikkeli* ‘article’ or *sika* ‘pig’ and *yrittäjä* ‘company boss’ within the same category, apart from a very abstract conceptual category THING. But pearls and pigs are obviously not more prototypical members of this category than articles and bosses. So prototypicality does not seem to constrain variation. As for implicational hierarchies, they have to include a well-defined set of universal categories. Non-default lexical constituents in Figure 84 above cannot be regarded as such universals.

As is shown in Chapter 5, restrictions on lexical variation are provided by the PU’s structure and the discourse context, e.g. the CS/PU or parts of it can serve as a basis for interaction between different constructions (Section 5.3.4); a syntactic constituent can be linked to non-default phonological structures of lexical items, which are semantically related to the default one (Section 5.3.3.2); non-default correspondence can be established between a constituent and phonological structure of its #SITUATION#/PU referent (Section 5.4.2).

Other variation one needs to account for in the *HELMIÄ SIOILLE* construction family concerns e.g. changes in case and number inflection of noun constituents like in Examples (126) and (127) below:

- (126) *Toinen näistä (Jukka) heitti viestissään vielä oikein helmen meille tyytymättömille sioille [...]*<sup>175</sup>

---

<sup>175</sup> GG: sfnet.tietoliikenne.yhteydentarjoajat/msg/a9beace372c5de66, Oct 3 2000

lit. 'One of these (Jukka) also threw in his message a real **pearl**<sub>SG GEN</sub> to us, dissatisfied pigs [...]'

(127) *Ei, en viitsi heittää helmiä **sialle** minäkään.*<sup>176</sup>

lit. 'No, neither do I bother to throw pearls to a **pig**<sub>SG ALL</sub>'

Obviously, additional levels are needed in the network in order to represent these morphosyntactic relations. As for the ability of the construction's constituents to appear in different morphological forms, construction grammarians would again suggest that the variable parts have to be schematized. For instance, if the verb in *KICK THE BUCKET* can be used in different tense-aspect-mood forms, then the idiom should be represented as *KICK*<sub>TNS</sub> *THE BUCKET* (Croft & Cruse 2004: 308-309). If one wants to account for examples like (126) and (127) above, the same principle has to be applied to the number and case marking of the noun constituents *helmi* 'pearl' and *sika* 'pig': *HEITTÄÄ*<sub>TNS</sub> *HELM*<sub>INUM CASE</sub> *SIK*<sub>ANUM CASE</sub>. However, the difference between more frequently occurring default forms and less frequent alternative forms will be eliminated as a result of such underspecification.

## 4.6 Summary

An overview and criticism of some main trends in idiom variation analysis were presented in this chapter. The central problem of the base form approach is that there are neither clear criteria of how such a form should be determined, nor adequate solutions for the methods of its representation that could account for possible variation. In the absence of such methods, classification principles of idiom variation lack any systematicity. The construction network model, on the other hand, enables one to step beyond the traditional approach to variation as manipulations with a postulated base form and an ad hoc classification of such manipulations as a primary means of analysis. We are now dealing with a PU's construction family instead of one construction.

However, by establishing a taxonomic hierarchy of constructions connected via more abstract levels one departs far from the original idea of family resemblance, where no construction is derived from the other, but instead constructions are interconnected via overlapping parts. As has been demonstrated above, schematic constructions are generalizations of whatever variable parts can be detected in the

<sup>176</sup> GG: sfnet.keskustelu.uskonto.kristinusko/msg/d86b88a778d780c3, Apr 9 1999

construction, but they neither specify any default form nor restrain variation. Dealing with variation in this way will easily lead to overgeneralization: not only will it allow for existing variation, but it will also license instances which happen to fit the same metaconstruction but cannot be related to the same idiom. One cannot simply assume that regular syntactic constructions will specify the same idiosyncrasies as substantive idioms do, so at some point of generalization all one achieves is an ad hoc syntactic and/or semantic categorization with no distinction between regular and construction-specific phenomena.

Unlike schematic syntactic configurations, idioms specify information practically at all levels of representation: phonological, morphological, syntactic and conceptual representations, referential indices and possibly some other cognitive levels of representation are linked together in their structure. Lexical and morphological items can themselves be considered mappings between conceptual, syntactic and phonological representations, which are decomposable to further levels (Nikanne 2005a). Since variation can simultaneously occur at several different levels, a detailed multi-level representation is needed in order to capture it. In other words, if one attempts to relate such big chunks of information as different instances of substantive idioms, one needs to go inside them and decompose them into smaller parts. A more detailed description will provide one with a more intricate system of linking than simple instance and subpart inheritance links, which are too abstract and can only capture subcategorization relations. Then we can abandon the description in terms of metaconstruction levels, since virtually nothing prevents one from relating instances of constructions directly via overlapping parts. The next chapter will be devoted to the further development of formal analysis of idiom variation in the framework of Conceptual Semantics and the Tiernet model.

## **5. The Tiernet approach to the analysis of variation in PUs**

### **5.1 Principles of the Tiernet-based model of the construction family network**

In Chapter 4, Section 4.5 I presented an overview of the inheritance hierarchy approach to the notion of construction family and variation. It has been criticized for a lack of descriptive power: since variation usually occurs simultaneously at several different levels, simple instance and subpart inheritance links between large “chunks” of information are not able to fully capture it. In Section 4.2 of the same chapter, some ad hoc categorizations of variation, e.g. into formal vs. semantic, usual vs. occasional were also critically reviewed. In this chapter I will demonstrate how an explicit conceptual-semantic model of an idiom’s structure (presented earlier in Chapter 3) can be applied to the description of different aspects of its variation. The major innovation and advantage of the Tiernet model is that it gives one the ability to abandon both the hierarchical models and variation classes. In what follows, I will bring forward some important principles, which the model is based upon.

#### **5.1.1 The notion of construction in the present approach**

In Section 3.3 of Chapter 3 I already pointed out that, apart from the distinction between regular syntactico-semantic mapping and irregular construction mapping, conceptual-semantic understanding of constructions (Nikanne 2005a) is akin to that of Goldberg’s (1995). I have also adopted Nikanne’s (2005a) definition of constructions as linking devices (correspondence rules between different levels of representation, or tiers) that license irregular linking patterns. In this respect both the default PU pattern and every single non-default PU realization are constructions. Together they form a vast network of what I here refer to as a PU’s construction family. However, the notion of family in the current approach differs from the inheritance model in that there is no hierarchical relationship between the patterns of the network. They are not mother and daughter constructions, but instead are all equal, and the system itself is dynamic and open to new linking possibilities. The present approach also differs from the base form vs. variant model,

which treats variants as deviations or derivatives from the base form. In the network model, I do not make a contrast between the base form and the derivate variants. Thus, the only difference is made between default and non-default linking<sup>177</sup> between different tiers, determined on the basis of quantitative analysis of empirical data. The linking pattern whose all links receive default status is considered to be the default construction.

In Section 4.4.2 I demonstrated that the default form of a PU is actually not the most usual form, but rather a combination of the most usual links that make up this form. In order to achieve a better comprehension of the concept of default one needs to describe variation on the micro-tier level. In principle, one could take any feature in the network and show variation around it. But, obviously, a separate detailed description of every idiosyncrasy within the family would make a very long list: variation occurs practically at all levels of representation (phonological, morphological, syntactic etc.), and since all these can occur simultaneously, the number of combinations would be very high. Merely listing all combinations or arranging them in some sort of classification does not necessarily help me to achieve the goal of this study, which is to gain understanding of how idioms behave in natural language. In order to understand the architecture and functioning of the whole system one needs to trace regularities in mapping between tiers both within and across modules. However, I will not commit myself on the matter of whether these patterns are constructions or not.

### 5.1.2 Constructions as networks

Since constructions are linking patterns between different levels of representation (Nikanne 2005a), each construction can be represented as a combination of links licensed by it. Previously in conceptual-semantic theory (i.e. in the works of Jackendoff, Nikanne, Pörn and Paulsen) the majority of correspondence linking has been formalized by means of letter or number indices. In the present study, in order to achieve a more explicit visualization, I have chosen to draw connecting dashed lines between corresponding units of the structure. In the most generalized form these links can be represented in a form of a triangle (Figure 85

---

<sup>177</sup> Linking types are introduced in Section 3.2.1 and linking values in Section 3.4.2 of Chapter 3.

below), with phonological, morphological and syntactic structures (PS-MS-SS), the conceptual structure (CS) and the referential structure (REF) at its vertices<sup>178</sup>. However, as was pointed out in Chapter 3, Section 3.4.3.2, in ambiguous phraseological units (i.e. syntactically well-formed units that can be assigned both literal and figurative interpretations), a single phonological and syntactic structure is mapped to the conceptual level both according to regular principles of syntactico-semantic correspondence linking (PS-MS-SS---CS) and according to idiosyncratic correspondence linking licensed by a PU (PS-MS-SS---CS/PU). The same double linking is repeated in the mapping of phonological and syntactic levels of representation to the referential tier (PS-MS-SS---#SITUATION# and PS-MS-SS---#SITUATION#/PU) (Figure 86 below). The idiosyncrasy of SS---CS/PU linking lies in the fact that it lacks a regular node-to-node mapping between the CS/PU's thematic arguments and the tier of syntactic categories via the tiers of DAs and syntactic functions, like in SS---CS. Instead, both structures correspond to each other as chunks. The only node-to-node correspondence that thematic arguments of the CS/PU are able to establish with other levels of representation comes from the referential tier. This happens in so-called semantically decomposable idioms, which I prefer to call idioms with #SITUATION#/PU referential constituents (Chapter 3, Section 3.4.3.4).

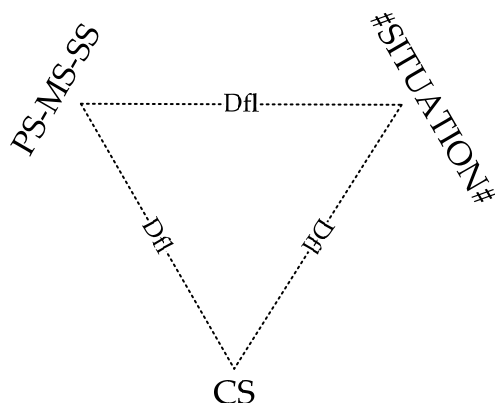


Figure 85 Unambiguous construction

<sup>178</sup> This basically corresponds to the classical linguistic or semiotic triangle (Ogden et al. 1923) where the vertices are the denotation (the object), the signifier (the symbol) and the signified (the concept).

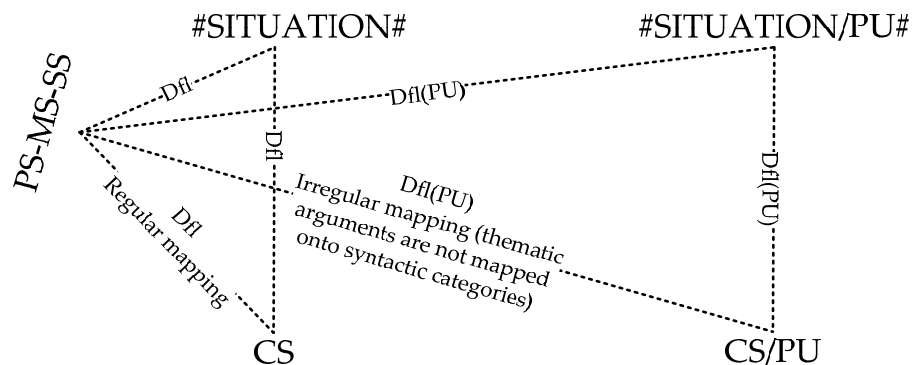


Figure 86 Ambiguous PU

Linking in the above Figure 85 and Figure 86 is, of course, largely oversimplified. Modules, which are positioned on the triangle's vertices, are comprised of multiple tiers with their own primitives, principles of their combination and interface rules. Some of these tiers have been discussed in more detail in Chapter 3 Section 3.2. This greatly complicates the picture: a truly detailed multi-tier representation of a construction results in a network where nodes are (directly or indirectly) units of different tiers. Figure 87 and Figure 88 below present some nodes which could be interconnected within a single construction.

The tripartite division into PS-MS-SS, CS and REF is preserved, but, as opposed to the above Figure 85 and Figure 86, vertices of the triangle are turned into its sides. In an even more detailed description a figure could, in fact, have more sides, i.e. instead of a triangle one would get a rectangle, a pentagon etc. This could be achieved by placing information of separate modules on their own side of the figure. However, in the present study for the sake of facilitation of conceptualization I have chosen to locate patterns which belong to syntactic, phonological and morphological modules (i.e. modules that are traditionally treated as form of a linguistic expression) on the same side of a triangle.



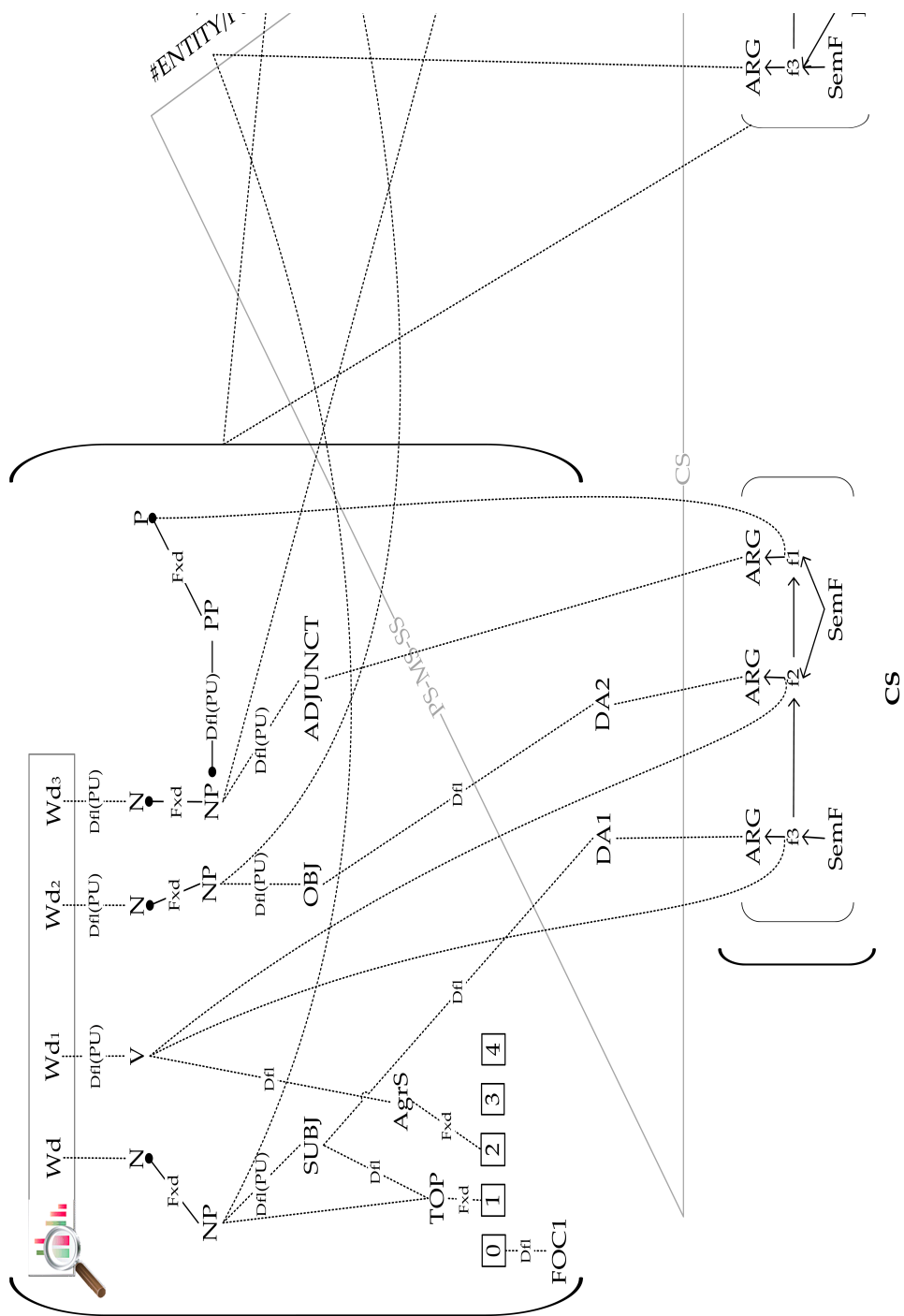


Figure 87 Nodes and links of a single construction

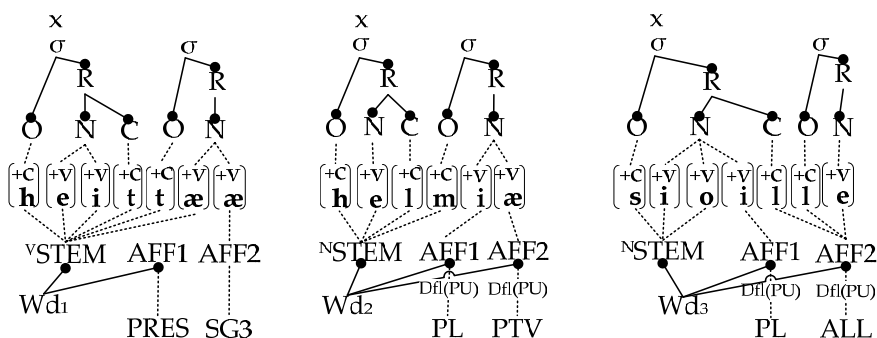


Figure 88 Nodes of morphosyntactic and morphophonological tiers

### 5.1.3 Variation as a construction family network

An important consequence of the Tiernet method of construction representation is that constructions themselves do not appear as nodes in a construction family network like they do in construction grammars. Instead, when several constructions are incorporated into a single network, each construction still remains a unique combination of nodes and links, as shown in Figure 87 above. Therefore, variations in the network are not considered as being derived from a base form, or some abstract metaconstruction, as has previously been done in different variation models. My model is in accord with Jackendoff's view of competence grammar, which is not formulated in terms of derivations, but in terms of constraints, or node admissibility conditions (McCawley 1968). Each constraint licenses a relation between two nodes of linguistic structure, and a structure is well-formed if it conforms to all applicable constraints. It is also important that constraints can be used to license structures at any point in the network (Culicover & Jackendoff 2005: 15). Thus, variability or the lack of such can be examined on an internode basis. In addition to general constraints of grammar, a PU can apply its own constraints that determine the extent to which its structure can vary. In Section 3.4.2 of Chapter 3 I have introduced the system of linking values (such as fixed, default or non-default), which can be used to formalize such constraints.

Similarities between constructions in a family can be described as overlapping nodes, while differences are always tier-specific. Figure 89 below presents an example of some possible constructions in the construction family of the Finnish PU *X HEITTÄÄ HELMIÄ SIOILLE* | *NP<sub>SUBJ</sub>[N{X}] Vtr{heittää 'throw'} NP<sub>OBJ</sub>[NP<sub>PL</sub>{helmi 'pearl'}] PP<sub>ALL</sub>[NP<sub>PL</sub>{sika*

'pig'} | lit. 'X throws pearls to pigs', id. 'X causes a transfer of some entity, evaluated by the speaker as good, to a recipient, evaluated as inadequate in some way' (henceforth referred to as the *HELMIÄ SIOILLE* construction family, or *HHS*). It shows overlapping nodes with a solid grey filling and construction-specific nodes with no filling. All constructions in Figure 89 share a similar syntactic structure, phonological form of the NP-object *helmiä* 'pearl<sub>PL</sub> PTV' and the PP-adjunct *sioille* 'pig<sub>PL</sub> ALL', their syntactic arguments are linked to the same argument positions on the thematic tier and thematic functions in the CS. Variation is observed in the phonological structure of the morphological word corresponding to the Verb constituent (*heittää* 'throw' vs. *antaa* 'give'), in the S-tier fields of the CS and the CS/PU, as well as in the referential tier. Linking between the SS and the CS depends on the lexical linking: the phonological form *heittää* 'throw' is associated with the CS, whose f-chain functions are selected by the S-tier Physical and Spatial fields; while the form *antaa* 'give' corresponds to the CS, which features the Possessive field in the S-tier. There is a similar dependence in correspondences between a particular #SITUATION#/PU on the referential tier and particular semantic fields in the CS/PU: the f-chain selected by the Communicative and the Cognitive/Perceptual fields corresponds to #COMMUNICATE#, and the f-chain selected by the Possessive field corresponds to #SELL#. Correspondence linking drawn from the PS-MS-SS to the #SITUATION#/PU and the CS/PU is presented as being independent of the variation in the lexical linking of the Verb constituent, i.e. two constructions featuring the same #SITUATION#/PU and correspondingly the same CS/PU can differ in the PS-MS-SS. These linking possibilities are summarized in Figure 90 below.

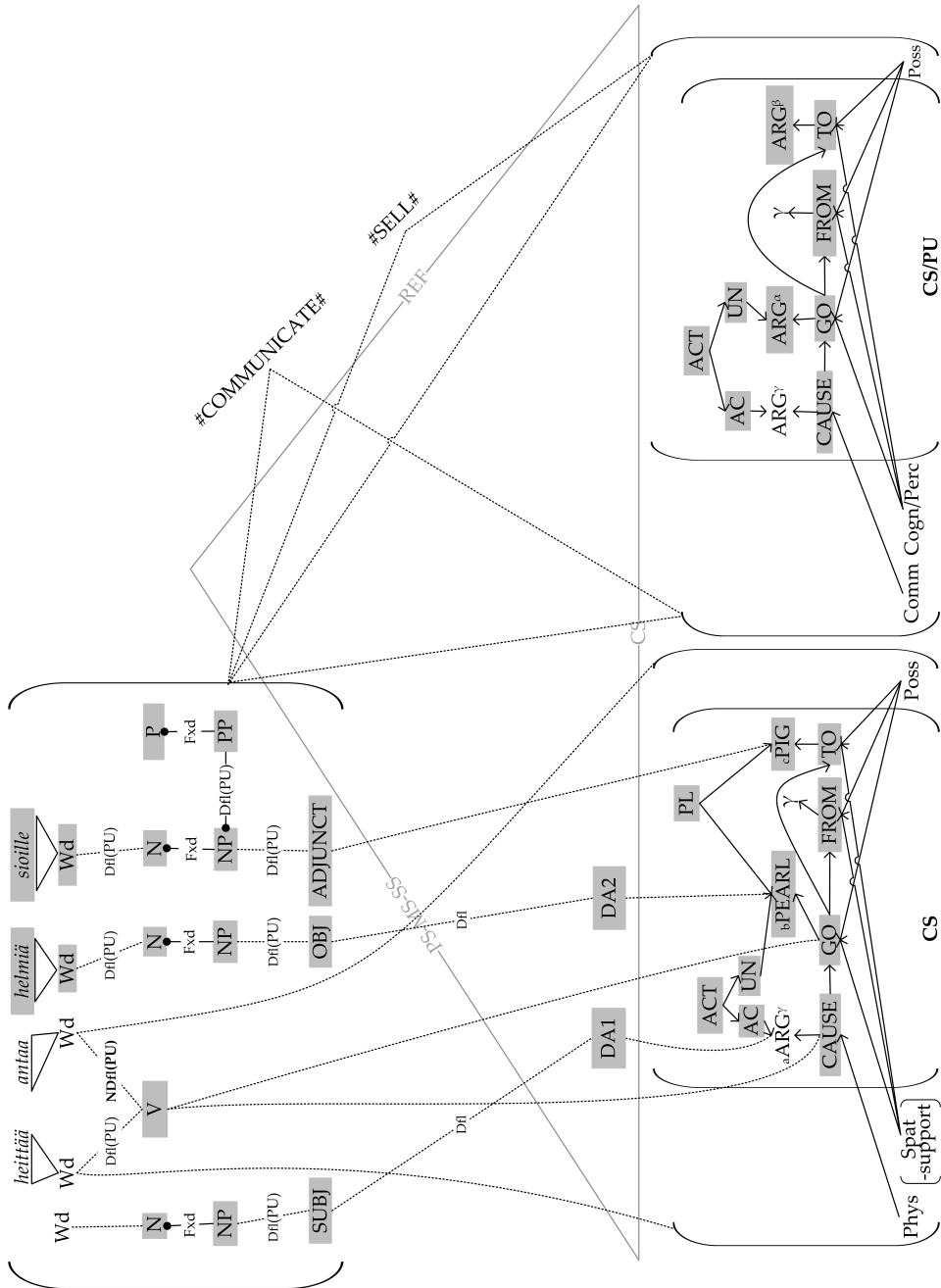


Figure 89 Overlaps and differences in some possible constructions

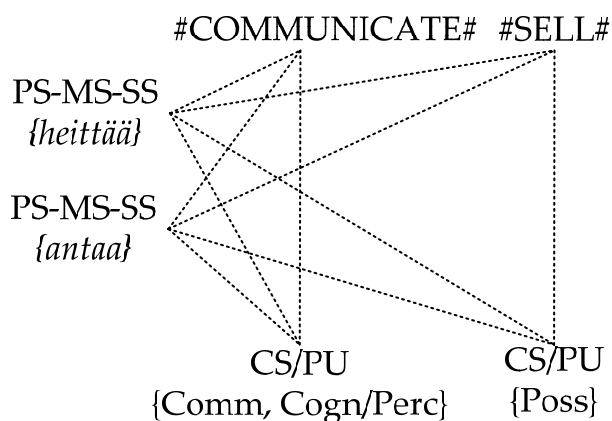


Figure 90 Linking possibilities between the PS-MS-SS, #SITUATION#/PU and the CS/PU in constructions from Figure 89

Figure 89 above presents just a very small portion of the whole construction family network of the idiom in question, which includes a substantial number of one-to-many pairings between the nodes on all three sides of the triangle. The complexity of the phenomenon can be better observed in Figure 91 below, which is an attempt to fit all S-tier nodes of the CS/PU and #Entity#/PU REFs observed in my data for the *HELMIÄ SIOILLE* construction family into a single description<sup>179</sup>. This linking will be examined in more detail in connection with the discourse-related variation in PU in Section 5.4.1 of the current chapter.

<sup>179</sup> The detailed syntactic analysis has been omitted for purely technical reasons.

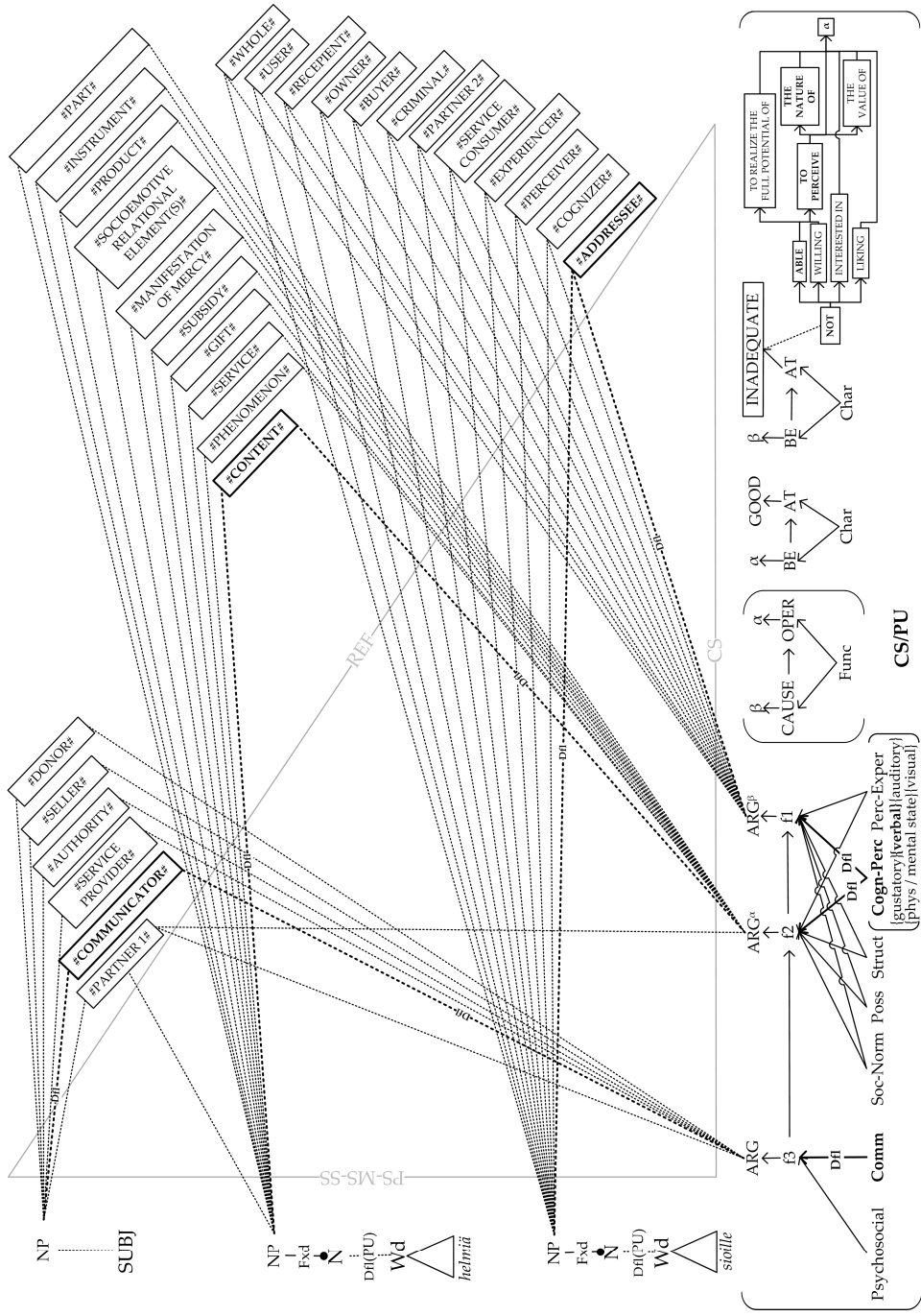


Figure 91 Variation in the CS/PU and #Entity#/PU linking

The question of whether every single unique combination of nodes and links has to be regarded as a construction of its own still remains open. The classic notion of a construction as a one form – one meaning pairing, which is reflected in e.g. Croft & Cruse's (2004: 263) idea that any morphological, syntactic, lexical, semantic, pragmatic or discourse-functional quirk is sufficient to constitute an independent construction, would suggest that, but it does not necessarily lead one to a better understanding of the phenomenon. Instead, one needs to be able to demonstrate how variation is licensed. In Section 5.2 I will examine internal and external cohesion, i.e. relations within the PU structure, on the one hand, and between this structure and the discourse context, on the other, as two underlying licensing sources of variation in PUs. The same differentiation will be followed in the subsequent Sections 5.2.3 and 5.4.

## **5.2 A PU's textual source, structure and discourse context as an underlying cause of variation**

In order to be able to explain what makes it possible to change the internal structure of a PU, even to the point when its single lexical constituents occur without immediate structural relation to each other, one has to look at the structural, contextual and intertextual properties of this PU. On the levels of syntax and phonology, PUs are multi-word phrasal units, with an inherent potential for structural variation. On the level of CS, PUs are semantic units, i.e. they are interpreted and treated as a single, coherent whole despite their structural separateness. On the level of discourse context, a PU refers to some situation and its entities. On the intertextual plane, a PU may have an identifiable source – a text fragment or passage. A PU's variability is therefore a result of the complex interplay between its phonological, morphological, syntactic and conceptual structures, context reference, and textual dependence. These can serve both as permissive and restricting factors of variation.

The ability of PUs to undergo considerable formal changes in discourse without loss of access to their CS/PU is secured by virtue of cohesion. According to Halliday & Hasan (1976) textual cohesion is a "semantic relation between an element in the text and some other element that is crucial to the interpretation of it". Phraseological cohesion has been earlier explored within the framework of applied stylistics by

Naciscione (2001: 51), who defines it as “part of the meaning of the base form; the unity of phraseological meaning in instantial use; a semantic and stylistic relation, realized in discourse by virtue of ties with the base components”. Naciscione (2001: 50) points out that phraseological cohesion provides a link between phraseological components in discourse, while components themselves are dependent on the PU and the context. She concludes that the totality of these dependencies results in cohesion.

As long as the structural and the textual dependencies involve different kinds of phenomena, they could be kept apart for methodological reasons. That is why I preferred in Petrova (2007b) to make a distinction between two types of phraseological cohesion, namely PU-internal, i.e. cohesive links between the constituents of the PU, and PU-external, i.e. cohesive links between the discourse context and the PU's structure. The autonomy of internal and external types of cohesion can be illustrated with Example (128) below, where *helmet* ‘pearl<sub>IPL NOM</sub>’ and *sioille* ‘pig<sub>PL ALL</sub>’ – lexical constituents of the Finnish PU *X HEITTÄÄ HELMIÄ SIOILLE* | *NP<sub>SUBJ</sub>[N{X}] Vtr{heittää ‘throw’} NP<sub>OBJ</sub>[N<sub>PL</sub>{helmi ‘pearl’}] PP<sub>ALL</sub>[NP<sub>PL</sub>{sika ‘pig’}]* | lit. ‘X throws pearls to pigs’, id. ‘X causes a transfer of some entity, evaluated by the speaker as good, to a recipient, evaluated as inadequate in some way’ – occur within the syntactic structure of the same clause:

(128) *Helmet eivät kelpaa ravinnoksi sioille.*<sup>180</sup>

‘Pearls are unfit for pigs’ nutrition.’

Taken outside of the context this utterance would allow two interpretations: firstly, the regular CS, which encodes a situation with actual pearls and actual pigs as thematic arguments and referents, and, secondly, since the two lexical items cohere with each other and the PU, the CS/PU reading accessibility is ensured by the PU internal cohesion, even though the default form is altered beyond recognition. This means that the CS/PU could be activated despite the absence of contextual support. In order to support the CS interpretation of *helmi* ‘pearl’ and *sika* ‘pig’ the context has to provide with explicit cohesive elements that indicate that one is talking about feeding real pearls to real pigs, as happens in Example (129) below:

<sup>180</sup> GG: sfnet.keskustelu.kielipolitiikka/msg/c855b46a87a90173, Mar 5 2003



- (129) *Pätkähullu miljonääri yritti syöttää helmiä sioille, koska uskoi että siat syövät helmiä. Hän ei ymmärtänyt että helmet eivät kelpaa ravinnoksi sioille.*  
 ‘A mad millionaire tried to **feed pearls to pigs**, because he believed that **pigs eat pearls**. He didn’t understand that **pearls** are not suitable as **nutrition** for **pigs**.’

In the same way the context can support the CS/PU reading with some elements that are semantically related to it, e.g. in Example (130) below the verb *perustella* ‘justify’ does not semantically cohere with the CS:

- (130) A: [...] kukaan ei osaa **perustella** asiallisesti suomalaisten systemaattista ja erittäin pahalle haisevaa pakkoruotsitusta. Et edes sinä [...]. Vai haluatko yrittää?  
 B: En. **Helmet** eivät kelpaa ravinnoksi **sioille**, ei vaikka kuinka kimaltelisivat.<sup>181</sup>  
 lit. ‘A: [...] no one can reasonably **justify** the systematic and extremely stinking enforced Swedification of Finns. Not even you [...]. Or do you want to try?’

B: No. **Pearls** are unfit for **pigs’** nutrition, no matter how they shine.’

In practice, however, it is quite difficult to distinguish between the internal and the external types of cohesion, due to their constant interplay in discourse, i.e. the information necessary to make the unambiguous interpretation is usually provided both by co-occurrence of constituents and the context.

### 5.2.1 Association by virtue of lexical co-occurrence

Phraseological units can be in principle regarded as collocational patterns where cohesion between lexical items is created by their regular co-occurrence within the same unit. Collocation is an ambiguous term which is used at least in three different senses.

In phraseology, the term collocation is used to denote a subclass of set phrases (semantic PUs). The meaning of a two-word collocation includes intact the meaning of one its constituents, but the other component of its meaning is expressed by an element contingent on the first. Mel’čuk (1998: 23) states that collocations constitute the absolute majority of PUs. The term restricted collocation is also used to describe the same subcategory by Cowie (1981), Gläser (1988) and Howarth (1996).

According to Halliday & Hasan (1976: 284), collocation is a subtype of lexical cohesion: “cohesion that is achieved through the association of lexical items that regularly co-occur”. Tanskanen (2006: 32, 35) points out

<sup>181</sup> GG: sfnet.keskustelu.kielipolitiikka/msg/c855b46a87a90173, Mar 5 2003

that the general definition of collocation in Halliday & Hasan (1976) is elusive and vague:

[...] but they do try to clarify it: the association is achieved when the lexical items have a tendency to appear in similar lexical environments or when they are related lexicosemantically (Tanskanen 2006: 33).

In Halliday (1985) collocation covers instances in which the relationship of the items depends on the association between them. Halliday (1985: 312-313) calls collocation a “co-occurrence tendency”, and states that “collocation is one of the factors on which we build our expectations of what is to come next”.

Collocation in corpus lexicography and lexical semantics refers to the lexical relation between two or more words which have a tendency to co-occur within a few words of each other in running text (Stubbs 2001: 24). A node-word co-occurs with collocates in a span of words to left and right. A node is the word-form or lemma being investigated. A collocate is a word form or lemma which co-occurs with a node in corpus. What is a node and what is a collocate depends on the focus of study. Definition of collocation is a statistical one: first of all it means frequent co-occurrence (Stubbs 2001: 29). As Tanskanen (2006: 33) remarks, cohesive collocation of Halliday & Hasan (1976) is not totally unrelated to the lexicographic collocation. The difference between these two is in the proximity of the items. In corpus lexicography, collocation refers to adjacent items: the number of collocates on either side of the node is typically restricted from four to six. Cohesive collocation refers to connections between longer stretches of a text (clauses and sentences). Therefore, if items occur next to each other, they are an instance of lexicographic collocation, but if they are separated by a longer stretch of text, their relationship can be regarded as an instance of cohesive collocation (Tanskanen 2006: 33-34).

The fact that lexical items form an associative link because they tend to co-occur, i.e. because they are related by a collocation or an idiom, can be demonstrated in the Russian language with the help of the Russian associative dictionary (RAD) (Karaulov 1994a, 1994b, 1996), which captures association-based speech patterns typical of modern language usage, labelled by the author of the dictionary by the term associative-verbal network. The dictionary based on data from numerous psycholinguistic tests, reflects int. al. the syntagmatic relations (i.e. linear co-occurrence) between words in conventional expressions and

phraseological units. In practice this means that words produced as associations to a given stimulus are associated to the stimulus by virtue of their mutual co-occurrence within the same conventional unit. For instance, the stimulus word *баран* [*barán*] ‘ram’ stimulated int. al. the following associations (Karaulov 1994b: 14):

- *ворота* [*voróta*] ‘gate’
- *на новые ворота* [*na nóvyje voróta*] ‘at a new gate’
- *на ворота* [*na voróta*] ‘at a gate’
- *новые ворота* [*nóvyje voróta*] ‘a new gate’
- *как на новые ворота* [*kak na nóvyje voróta*] ‘like at a new gate’
- *у ворот* [*u vorót*] ‘near a gate’
- *у новых ворот* [*u nóvyh vorót*] ‘near a new gate’
- *уставился как баран* [*ustávils’a kak barán*] ‘stares like a ram’

All these associations reflect the fact that *баран* [*barán*] ‘ram’ is a lexical constituent in a Russian PU *X СМОТРИТ/УСТАВИЛСЯ НА Y, КАК БАРАН НА НОВЫЕ ВОРОТА* [*X smótrit/ustávilsja na Y, kak barán na nóvyje voróta*] | *NP<sub>SUBJ</sub>[N{X}] V{смотреть [smotrét’] ‘look’/уставиться [ustávit’s’a] ‘stare’} <PP[P{на [na] ‘at’} NP{Y}]> CMPR C{как [kak] ‘like’} NP{баран [barán] ‘ram’} PP[P{на [na] ‘at’} NP<sub>PL</sub>[AP{новый [nóvyj] ‘new’} N{ворота [voróta] ‘gate’}]]* | lit. ‘X looks/stares <at Y> like a ram at a new gate’, id. ‘X stares <at Y> dumbly’. There is also evidence which speaks in favour of the existence of associative links between the meaning of the idiom and its constituents: e.g. the stimulus word *безделье* [*bezdel’je*] ‘idleness’ produced the idiom *X БЬЁТ БАКЛУШИ* [*X b’jot baklushi*] | *NP<sub>SUBJ</sub>[N{X}] V{бить ‘break, smash’} NP<sub>OBJ</sub>{баклуши<sup>182</sup> [baklúši] ‘small chunks of wood chopped from large blocks, blanks for woodwork’}* | lit. ‘X chops blanks for woodwork’, id. ‘X is idle, does nothing, twiddles his/her thumbs; fritters away/waste his/her time’ (Karaulov 1996: 15) and the stimulus *бездельник* [*bezdel’nik*] ‘idler’ was associated with the single constituent of this idiom – the word *баклуши* [*baklúši*] (Karaulov 1994a: 9). Karaulov (ibid. 192) also observes that the grammatical form of the stimuli influenced the make-up and grammaticalization of reactions in the respective associative field, which can in principle lead to stronger associative links between lexical constituents of the PU if they are presented in their default morphological form. The number of PUs in the associative-verbal network turned out to be quite considerable: on estimate PUs occur altogether 17 thousand

<sup>182</sup> An obsolete, cranberry lexical item, which nowadays occurs only in this PU.

times and the number of different PUs exceeds 2.5 thousand. The author made an observation that informants did not use the full form of a PU in their answers, but instead gave an abbreviated, condensed and sometimes transformed form, which in the author's opinion can correspond to the form in which the PU is stored in the memory of a language speaker. These facts also nicely contradict with the approach which equates the base-form of a PU with its dictionary form and points towards the more analytic model of PU representation.

A similar pilot test, which I have performed with the Edinburgh Associative Thesaurus (EAT)<sup>183</sup> provides some empirical association data for the English language and also demonstrates an interesting asymmetry in associative links. For example, the word *bucket* stimulated int. al. *kick* as its association, which is explained by their co-occurrence in the idiom *KICK THE BUCKET*. However, the verb *kick* did not produce associations with *bucket*. The stimulus *leg* produces an association to *pull* (i.e. *PULL SOMEBODY'S LEG*) but not vice versa. *Dust* is listed as an association of the stimulus *bite* (i.e. *BITE THE DUST*), but not vice versa. There are also items that stimulate associations in both ways, e.g. in pairs like *coffin* and *nail* (*A NAIL IN SOMEBODY'S COFFIN*), *grapes* and *sour* (*SOUR GRAPES*), *rags* and *riches* (*FROM RAGS TO RICHES*) and *needle* and *haystack* (*LOOK FOR A NEEDLE IN A HAYSTACK*) both items have each other as associations. EAT also shows that constituents tend to associate better if they appear in their default morphological forms, e.g. the proportion of occurrence of the lemma *riches* produced as response to the stimulus *rag* in singular form was 0.01, compared to the proportion of 0.24 stimulated by the plural form *rags*. The proportion of occurrence is the frequency of a particular response divided by the total count of responses to the stimulus word (rounded to two decimal places). The proportion of occurrence can be multiplied by 100 to provide a percentage; i.e. in this case 24% of respondents produced *RICHES* as the response to the word form *RAGS* and 1% produced it as the response to the word form *RAG*<sup>184</sup>.

---

<sup>183</sup> The number of subjects used for data collection in EAT is 100, which is considerably smaller than that in RAD, where about 11000 subjects were used.

<sup>184</sup> Čermák (2007: 21) actually goes as far as suggesting that idioms are in fact made up of word forms and not lexemes or words.

### 5.2.2 Reference to discourse context

Cohesion presupposes that the interpretation of a particular element of the text is dependent on that of another, i.e. the one cannot be effectively decoded except by recourse to the other (Halliday & Hasan 1976: 4). Since the interpretation of a PU token is secured by its semantic relation with other elements of its local linguistic context, PU external cohesion can be defined as a semantic relation between the discourse context and the PU's internal structure. In other words, external cohesion is created by elements of the discourse context co-referential with the elements of the PU's conceptual structure.

Since contextual elements can be categorized in many different ways, the notion of context used in literature is a hypernym, which, depending on the research goals, covers a variety of multifaceted phenomena, both linguistic and non-linguistic. Therefore, it is necessary to specify what I mean by it in the scope of this study and to narrow down my technical characterization of context in accordance with the data and the object of analysis. Firstly, inasmuch as my entire data is derived from the CMC (Computer-mediated Communication), the PU tokens appear in a local situational context of a certain virtual community (Rheingold 2000), i.e. a group of individuals who use networked computers in order to interact through specific media (in my case Usenet newsgroups<sup>185</sup>) for social or other purposes. However, I am not able to include such social or situational aspects as age, gender, location or relative status of discourse participants in my notion of the context, since these variables are largely unknown to me. Secondly, the context is not necessarily limited to the immediate textual surrounding (local linguistic context) of a PU, but can also include knowledge shared by interlocutors about the wider linguistic context (e.g. previous experience of other texts on discussion groups that can be crucial for the interpretation of a given text). However, this kind of intertextual knowledge is not immediately available to a detached observer (except for the cases where the message contains quotations from previous messages in the thread), who most of the time will have to rely solely on the local linguistic context of a PU. Therefore, by the PU's context I understand first of all the local linguistic context of a PU, which includes the analyzed unit as well as elements that appear in a

---

<sup>185</sup> Usenet newsgroups as a source of idiom variation data have been described in Section 1.3 of the Introduction.

certain structural (cohesive) relation with elements of the unit's conceptual structure and that are relevant for its interpretation.

Prior turns in the discussion that appear in the form of quotations in the body of the same message before the PU in focus can also be regarded as its local linguistic context. The same goes for the topic of the thread and the name of the discussion group, which also can provide information relevant for interpreting the PU. For instance, in the following example presented in Table 6 below the subject area of the discussion group (sports, football) determines thematic boundaries for all its discussions within this area, while the subject of the first discussion narrows it specifically to Premier League.

Table 6 Example of reference to wider linguistic context in a Usenet newsgroup

Group	Discussion/Date	Writer	Message text
[sfnet.urheilu.jalkapallo 'sfnet.sports.football'] <sub>3</sub>	Valioliiga- sovinismia 'Premier League chauvinism'/ [Nov 20,22 1998] <sub>2</sub>	Erkki D	<i>Kanu kanuunoihin, kiva.</i> <sup>186</sup> 'Kanu to cannons, cool.'
		[Marko Pykäläniemi] <sub>1</sub>	<i>Toivottavasti. [Kankeat koikkelehtijat ovat aina tervetulleita Valioliigaan, kunhan ne vain siirtyvät muualle kuin Manuun]<sub>4</sub>...</i> <sup>187</sup> 'I hope so. [Clumsy wobblers are always welcome to Premier League as long as they transfer elsewhere than to Manu] <sub>4</sub> .'
		Erkki D	<i>Kyllä, ei helmiä sioille.</i> <sup>188</sup> 'Yes, no pearls to pigs.'

<sup>186</sup> GG: sfnet.urheilu.jalkapallo/msg/11353ea3999000ce, Nov 20 1998

<sup>187</sup> GG: sfnet.urheilu.jalkapallo/msg/281ee0c618f5a16c, Nov 22 1998

<sup>188</sup> GG: sfnet.urheilu.jalkapallo/msg/14c6875e1dc8ac62, Nov 22 1998

	<p><i>Tilanne!</i>  ‘Standing!’/  [May 18 1999]<sup>2</sup></p>	<p><b>Erkki D</b></p>	<p><i>Mitä tulee [jonkun valopään]<sub>1</sub> [aikoinaan]<sub>2</sub> [ryhmässä]<sub>3</sub> päästämään [sammakkoon Nwankwo Kanusta "koikkelehtijana, joka ei olisi tervetullut ManUun"]<sub>4</sub>, niin kyllä, ei todellakaan saa <b>antaa helmiä sioille</b>. Kauden ylivoimaisesti hinta-laatusuhteeltaan paras hankinta Valioliigassa.<sup>189</sup></i>  ‘As to [the gaffe]<sub>4</sub> [once]<sub>2</sub> made by [some bright spark]<sub>1</sub> in [the group]<sub>3</sub>, [about Nwankwo Kanu as “a wobbler, who wouldn’t be welcome to ManU”]<sub>4</sub>, it is true, indeed one mustn’t <b>give pearls to pigs</b>. The season’s best acquisition in the Premier League from the point of view of price-quality ratio.’</p>
--	---	-----------------------	---

Here I am looking at four separate messages posted in two discussion threads of the same newsgroup. In the first message of the first discussion thread (*Valioliiga-sovinismia* ‘Premier League chauvinism’), the word *kanuunat* ‘cannons’ refers to the Arsenal Football Club, which has a picture of a cannon on its official emblem. The remark as a whole is a comment on the fact that in the near future the football player *Nwankwo Kanu* is going to be signed by Arsenal. The message which appears 5 months later in the second discussion thread (*Tilanne!* ‘Standing!’) contains several references to the wider linguistic context, i.e. to the second message posted in the first discussion thread. Both this message and a message in the preceding conversation feature constructions of the same PU family – *EI HELMIÄ SIOILLE* | *NEG*{*ei* ‘not’} *NP<sub>OBJ</sub>*[*N<sub>PL</sub>*{*helmi* ‘pearl’}] *PP<sub>ALL</sub>*[*N<sub>PL</sub>*{*sika* ‘pig’}] | lit. ‘no pearls to pigs’ and *X ANTAA HELMIÄ SIOILLE* | *NP<sub>SUBJ</sub>*[*N*{*X*}] *V*{*antaa* ‘give’} *NP<sub>OBJ</sub>*[*N<sub>PL</sub>*{*helmi* ‘pearl’}] *PP<sub>ALL</sub>*[*N<sub>PL</sub>*{*sika* ‘pig’}] | lit. ‘X gives pearls to pigs’, which both refer to the same situation: *Nwankwo Kanu* was not invited to *Manchester United* (*ManU*); both constructions are used by the same writer (*Erkki D*). The second writer’s (*Marko Pykäläniemi*) comment in the first conversation contains a negative evaluation of *Kanu* as a ‘clumsy wobbler’ who, in his opinion, was not worthy of being accepted to *ManU* as a player, while *Erkki D* reacted by

<sup>189</sup> GG: sfnet.urheilujalkapallo/msg/238f22261c1a9a62, May 18 1999

using the construction *EI HELMIÄ SIOILLE | NEG{ei 'not'} NP<sub>OBJ</sub>[N<sub>PL</sub>{helmi 'pearl'}] PP<sub>ALL</sub>[N<sub>PL</sub>{sika 'pig'}]* | lit. 'no pearls to pigs'. Besides that the same PU is used in these two chronologically distant fragments, cohesion between them is established by anaphoric references which are made by *Erkki D* in his second message:

- The person-deictic element *joku valopää* 'some bright spark' refers to *Marko Pykäläniemi*.
- The time-deictic reference *aikoinaan* 'once, at one time' indicates that the referred fragment of discourse occurred in the past.
- The discourse-deictic element *ryhmässä* 'in the group' points to the fact that the referred fragment can be found within the limits of the same newsgroup, i.e. *sfnet.urheilu.jalkapallo* 'sfnet.sports.football'.
- *Marko Pykäläniemi*'s comment on *Kanu* from the first fragment is referred to as *sammakko* 'the gaffe' and is paraphrased.

The local context in the first conversation was not sufficient enough to establish *Erkki D*'s attitude towards the participants of the situation, and thus the construction *EI HELMIÄ SIOILLE | NEG{ei 'not'} NP<sub>OBJ</sub>[N<sub>PL</sub>{helmi 'pearl'}] PP<sub>ALL</sub>[N<sub>PL</sub>{sika 'pig'}]* | lit. 'no pearls to pigs' used by *Erkki D* could be interpreted either as assigning the argument referent role of *helmiä* 'pearls' to *ManU* and of *siat* 'pigs' to *Kanu*, or vice versa. The positive evaluation of *Kanu* which *Erkki D* presents in the second fragment, together with the construction *X ANTAA HELMIÄ SIOILLE | NP<sub>SUBJ</sub>[N{X}] V{antaa 'give'} NP<sub>OBJ</sub>[N<sub>PL</sub>{helmi 'pearl'}] PP<sub>ALL</sub>[N<sub>PL</sub>{sika 'pig'}]* | lit. 'X gives pearls to pigs', assigns the role of pearls to *Kanu* and thus clarifies the meanings of both the first and the second constructions.

The above example demonstrates that both the local and the wider linguistic context play an important role in the successful interpretation of an ambiguous PU. The topic of the discussion group provides a wider socio-cultural context, i.e. a broader background against which communication is interpreted (Hewings & Hewings 2005). This includes the football realities, the knowledge of which is expected from the discourse participants of the *sfnet.urheilu.jalkapallo* 'sfnet.sports.football' group, but which may be unknown to the researcher.

Thus, a proper description of semantic cohesion in the text is necessary in order to account for a PU's contextual meaning as well as recognition and interpretation of its modifications. I need a model of formal semantic analysis that could explicate semantic cohesion in the text, i.e. a model that could identify and overtly relate the parts of the text that are relevant



for the interpretation of the PU to parts of its semantic structure. The starting point of my analysis is an explicit description of the PU's internal structure, i.e. the conceptual-semantic Tiernet model of formal representation of PUs described in Chapter 3 and Section 5.1 of the current chapter. My claim is that the cohesive relationship between the fragments of discourse and a PU's structure can be explicated via its linking to the referential tier (discussed in Chapter 3, Section 3.4.3.4). In Section 3.4.3.5 of Chapter 3 I have already demonstrated how the formal description of reference in the PU *HEITTÄÄ HELMIÄ SIOILLE* 'throw pearls to pigs' can be applied to actual usage events: in Examples (80), (81) and (82) parts of the context were co-indexed with referential indices in Figure 65, Figure 66 and Figure 67, respectively. In Sections 5.2.3 and 5.4, I will attempt with the help of the Tiernet model to demonstrate explicitly to what extent variation is determined by structural properties of the PU, on the one hand, and the context of its realization, on the other.

### 5.2.3 Textual dependence – relation between a PU and its source

Dobrovolskij & Piirainen (2005: 103-103, 230f) discuss intertextual phenomena in conventional figurative units (CFUs) as a subtype of culture-based knowledge and use the term textual dependence to describe the "intertextual relation between CFUs and texts that can be identified as their sources". They also distinguish between two groups of such dependence – quotations and allusions. Quotations are direct references to particular written passages of texts, more or less word-to-word from fictional literature, the Bible, mass media etc. (e.g. the Finnish PUs in (131) and (132) below), while allusions are references to an entire text or a large section of it (e.g. the Finnish PUs in (133) and (134) below).

- (131) *EROTTAA JYVÄT AKANOISTA* lit. 'to separate the wheat from the chaff', id. 'to separate good, valuable entities from worthless ones' (textual source – the Bible, Matthew 3:17);
- (132) *SUOMUKSET PUTOAVAT JONKUN SILMILTÄ* lit. 'the scales fall from someone's eyes', id. 'someone begins to see things as they are' (textual source – the Bible, Acts 9:18);
- (133) *TAISTELU TUULIMYLLYJÄ VASTAAN* lit. 'a battle against windmills', id. 'a futile, absurd fight against some imaginary things' (textual source – Cervantes' *DON QUIXOTE*);
- (134) *AATAMIN PUVUSSA* lit. 'in Adam's costume', id. 'naked' (textual source – the Bible, Genesis 2).

Knowledge of a PU's actual textual source can be expressed by a speaker e.g. by using indirect speech (135), hearsay evidential (136) or quotative evidential (137), (138):

- (135) *Jeesus Nasaretilainen taisi sanoa saman asian jotenkin, että --- hmmm -- ainakin hän sanoi, ettemme saisi heittää sioille helmiämme.*<sup>190</sup>  
lit. 'I think Jesus of Nazareth said the same thing somehow, that --- hmmm – at least he said that we shouldn't throw our pearls to pigs.'
- (136) *Muistathan, miten muuan nuorukainen julisti: Älkä heittäkö helmiä sioille!*<sup>191</sup>  
lit. 'Do you remember how a certain young man preached: Don't throw pearls to pigs!'
- (137) *Se taitaa kuitenkin usein olla sitä "helmien heittämistä sioille", josta Jeesus varoitti (Matt. 7:6).*  
lit. 'However I think it is often a case of "throwing of pearls to pigs", which Jesus warned about (Matt. 7:6).'
- (138) *Raamattu sanoo "ei helmiä sioille".*<sup>192</sup>  
lit. 'The Bible says "no pearls to pigs".'

In most cases the average speaker is not conscious of intertextual phenomena, in which case the PU can be either ascribed to an unknown source (139), general folk wisdom (140), or to someone else than Jesus Virhe. Viitteen lähdettä ei löytynyt.:

- (139) *Helmiä sioille. - Tuntematon*<sup>193</sup>  
lit. 'Pearls to pigs. - Unknown'
- (140) *... kirjastojen ylläpito verovaroin on slummiutuneilla lähiöalueilla täysin turhaa rahan haaskuuta - "helmiä sioille" kuten vanha kansa sanoo.*<sup>194</sup>  
lit. '... the maintenance of libraries with tax revenue in decayed suburbs is a completely useless waste of money - "pearls to pigs" as the old folk say.'
- (141) *Äidinkielenopettajani vertauskuvallisin sanoin: ei helmiä sioille.*  
lit. 'In my Finnish language teacher's figurative words: no pearls to pigs.'<sup>195</sup>

Dobrovol'skij & Piirainen (2005: 81; 102) admit that knowledge of the textual source is not an obligatory condition for using and processing a PU and therefore it is of no importance whether or not speakers using a

<sup>190</sup> GG: finet.evl.keskustelu/msg/89bd0bbb59e2f78e, Dec 22 1997

<sup>191</sup> GG: sfnet.keskustelu.yhteiskunta/msg/ab1b805eef749d7e, Mar 21 2002

<sup>192</sup> GG: sfnet.keskustelu.evolutio/msg/ecab51bae772f974, Jun 19 1999

<sup>193</sup> GG: sfnet.keskustelu.yhteiskunta/msg/70eb3f107fd74482, Sep 15 2004

<sup>194</sup> GG: sfnet.keskustelu.politiikka/msg/dfb93b401f7283a3, Apr 23 2002

<sup>195</sup> GG: sfnet.harrastus.mp/msg/01764fec52579051, Aug 22 2002

PU are aware of its origin. However, on one occasion they indicate that this knowledge “is implicitly present in the plane of content” (ibid. 233). I do not share this point of view and thus do not include any knowledge of the PU’s source in the network model of PU structure.

Dobrovol’skij & Piirainen (2005) look at intertextual phenomena predominantly from the viewpoint of idiom motivation and certain structural idiosyncrasies. For instance, the motivation of the Finnish idiom *SIINÄ ON VILLAKOIRAN YDIN* lit. ‘(in this) there is the poodle’s core’, id. ‘that is what is behind it’ cannot be explained on the basis of the concept POODLE. However, according to the authors, a speaker’s knowledge of the fact that it is a quotation from Goethe’s drama FAUST could provide motivational links between this PU’s regular CS, on the one hand, and CS/PU, on the other. Since I am mostly interested in PU variation, in what follows I will briefly examine a few actual usage examples of the Finnish PU *HEITTÄÄ HELMIÄ SIOILLE* ‘throw pearls to pigs’, where the biblical context of this idiom<sup>196</sup> affects the context of its use and possibly acts as an underlying source of variation.

Both examples in (142) and (143) below contain the construction *SIAT TALLAAVAT HELMET JALKOIHINSA* | *NP<sub>SUBJ</sub>[N<sub>PL</sub>{sika ‘pig’}] V{tallata ‘trample’} NP<sub>OBJ</sub>[N<sub>PL</sub>{helmi ‘pearl’}] PP<sub>ILL</sub>[N<sub>PL3</sub>{jalka ‘leg, foot’}]* | lit. ‘pigs trample pearls under their feet’, which is a paraphrase of the original biblical passage. Strictly speaking, examples like (143) below present borderline cases, which could be regarded rather as biblical quotations than altered PU tokens. The source-indicating evidential *Raamatun mukaan* ‘according to the Bible’ and quotation marks emphasize the reported speech. However, this is not an exact quotation; it also contains the clause *sillä siat eivät ymmärrä niiden arvoa* ‘for pigs do not understand their value’, which is missing from the original.

- (142) *Mutta valitettavasti enemmistö lukijoista on sikoja jotka tallaavat aidot helmet jalkoihinsa* rynnistäessään mediapuffattujen mitättömyyksien perässä.<sup>197</sup>  
lit. ‘But unfortunately the majority of readers are pigs who **trample genuine pearls under their feet** storming after nobodies puffed up by media.’
- (143) *Kristinusko ryhmään on kirjoitettu suurimpia herjoja, mitä on nähty. Raamatun mukaan; “Älkää heittäkö helmiä sijoille, sillä siat eivät ymmärrä niiden arvoa, vaan tallaavat ne jalkoihinsa”*.<sup>198</sup>

<sup>196</sup> See Section 4.4.4 in Chapter 4 for more details on the original context.

<sup>197</sup> GG: sfnet.harrastus.kulttuuri.sarjakuvat/msg/1e9dad4f4f6f4921, Jan 16 1998

lit. 'The biggest insults ever seen have been written in the Christianity (news)group. According to the Bible; "Do not throw pearls to pigs, for pigs do not understand their value, but **trample them under their feet**".'

In (144) below the verb *tallata* 'trample' is substituted by a synonym verb *polkea* 'trample':

- (144) *Näin me siis opimme lisää, että viisaat eivät milloinkaan vaadi arvoa itselleen, koska annettu arvo on kuin kaivoon kannettua vettä, tai se on kuin kallisarvoinen helmi, joka heitetään **sikojen poljettavaksi**.*<sup>199</sup>

lit. 'Thus we learn more that wise men never require merit for themselves, because given merit is like water which is carried into a well, or it is like a precious pearl, which is thrown **to be trampled on by pigs**.'

In (145) below it is not the *pearls* that get trampled on by the *pigs*, by those who 'show pearls to pigs'. The verb *tallata* 'trample' is substituted by a derivative frequentative verb *talloa* 'trample repeatedly, stomp':

- (145) *Kirkossa ja seurakunnissa on kahdenlaista porukkaa. Niitä kenellä on helmiä, jotka tietävät olla antamatta sioille mitään. Ja niitä jotka ryöpyttävät sioille kaiken aikaa jotain, mutta takuulla eivät helmiä. Sitten on ne onnettomat väliinpuutoajat jotka **tulevat sikojen tallomiksi** esitettyään helmensä väärässä paikassa.*<sup>200</sup>

lit. 'There are two kinds of folk in the church and in congregations. Those who have pearls, who understand not to give anything to pigs. And those who whirl something to pigs all the time, but definitely not pearls. Then there are those unlucky losers who **get repeatedly trampled on by pigs** after having shown their pearls in the wrong place.'

Example (146) below contains several biblical quotations:

- (146) A: *Tuo on \_todella\_ lapsellista.*<sup>201</sup>  
 B: ***Tulkaa lapsen kaltaisiksi.*** :-)<sup>202</sup>  
 C: *Jeesushan sanoi näin.? Miksiköhän? Lapsiahan on helppo johdatella ja uskovat aikamoisia juttuja. Tulipa vain mieleen... No se siitä*<sup>203</sup>  
 B: *Entisaikaan tuohon olisi vastattu, että **ei helmiä sioille** eikä **kyykäärmeen sikiöille** joiden käsissä kaikki pyhä muuttuu raskaaksi kuin lyijy.*<sup>204</sup>

<sup>198</sup> GG: sfnet.keskustelu.uskonnottomuus/msg/0a33eada42cf1bf5, Mar 8 2000

<sup>199</sup> GG: sfnet.keskustelu.politiikka/msg/7ae85add16c48cdc, Nov 19 2002

<sup>200</sup> GG: sfnet.keskustelu.uskonto.kristinusko/msg/3b2f3302ce035c93, Oct 1 2001

<sup>201</sup> GG: sfnet.keskustelu.rajatieteet/msg/3e53816dc049fb75, Mar 23 2003

<sup>202</sup> GG: sfnet.keskustelu.rajatieteet/msg/7d76372b11388586, Mar 23 2003

<sup>203</sup> GG: sfnet.keskustelu.rajatieteet/msg/a51a089cd5440d47, Mar 23 2003

<sup>204</sup> GG: sfnet.keskustelu.rajatieteet/msg/3cdc9864d7eb53e5, Mar 23 2003

C: Niinpä. en olisi helmiä huolinutkaan sillä "**helpompi on kamelin kulkea neulansilmän lävitse kuin rikkaan päästä taivaaseen**"<sup>205</sup>

'A: That is \_really\_ childish.

B: **Become like children.** :-)

C: It was Jesus who said so? I wonder why? Indeed, children are suggestible and believe in old wives' tales. Just crossed my mind... Well that's it.

B: In times past an answer to that would be that **no pearls to pigs** nor to **brood of a viper** in whose hands everything sacred becomes heavy as lead.

C: Sure. I would not want pearls anyway for "**It is easier for a camel to go through the eye of a needle than for a rich man to enter the heaven**"

The first quotation is *Tulkaa lapsen kaltaiseksi* 'Become like children', which refers to the following Biblical passage:

(147) *Totisesti: ellette käänny ja tule lasten kaltaiseksi, te ette pääse taivasten valtakuntaan.*<sup>206</sup>

'I tell you the truth, unless you change and **become like little children**, you will never enter the kingdom of heaven.'

The second one is *kyykäärmeen sikiöille* 'brood/spawn of a viper' – the most interesting here, since it is connected by the coordinating conjunction *eikä* 'nor' to the construction *EI HELMIÄ SIOILLE* | *NEG{ei 'not'} NP<sub>OBJ</sub>[N<sub>PL</sub>{helmi 'pearl'}] PP<sub>ALL</sub>[N<sub>PL</sub>{sika 'pig'}]* | lit. 'no pearls to pigs'. This slightly modified expression comes from the Biblical passage presented in (148) below:

(148) *Mutta kun Johannes näki, että hänen kasteelleen oli tulossa myös monia fariseuksia ja saddukeuksia, hän sanoi heille: "Te käärmeen sikiöt! Kuka teille on sanonut, että te voitte välttää tulevan vihan?"*<sup>207</sup>

'But when John saw many of the Pharisees and Sadducees coming to his baptism, he said to them: "You **spawn of a viper**! Who told you that you can avoid the wrath to come?"

The third biblical quotation is a witty response to the PU *EI HELMIÄ SIOILLE* 'no pearls to pigs': the author of (C), who is the referent of *NP<sub>PL</sub>{sika 'pig'}*, remarks that he would not want 'pearls' anyway, because the Bible warns against riches in the following passage (149):

(149) *Helpompi on kamelin mennä neulansilmästä kuin rikkaan päästä Jumalan valtakuntaan.*<sup>208</sup>

<sup>205</sup> GG: sfnet.keskustelu.rajatieteet/msg/c8310f81868b8681, Mar 23 2003

<sup>206</sup> Matt. 18: 3

<sup>207</sup> Matt. 3: 7

<sup>208</sup> Mark 10: 25

‘It is easier for a camel to go through the eye of a needle than for a rich man to enter the kingdom of God.’

Example (150) below features the construction *X HEITTÄÄ HELMIÄ SIOILLE* | *NP<sub>SUBJ</sub>[N{X}] V{heittää ‘throw’} NP<sub>OBJ</sub>[N<sub>PL</sub>{helmi ‘pearl’}] PP<sub>ALL</sub>[N<sub>PL</sub>{sika ‘pig’}]* | lit. ‘X throws pearls to pigs’ in the first message and several allusions in the two following messages:

(150) A: *Voisitko tällä kertaa uhrautua kahden sanan selityksen sijasta monisanaisempaan ilmaisuun? Vai onko tuo taas sama kuin ”heittäisi helmiä sioille”? [...] Niin, olisi sama kuin heittäisi helmiä sioille. Ymmärrän.*<sup>209</sup>

B: *Minäpä pistän kärsäni tähänkin keskusteluun ja nappaan helmen niinkuin käärme Moosekselta. [...] SKRUNTS, OIIIINK! Sakari (Sikaileva telelogisti)*<sup>210</sup>

C: *Ja taas meni pieleen! Eihän se kärmes Moosekselta mitään helmeä napannut, vaan Mooses korotti sian erämaassa. Se kuuluisa helmi kuului Adalminalle.*<sup>211</sup>

lit. ‘A: Could you this time make an effort to give a more elaborate expression instead of a two-word explanation? Or is that again the same as “throwing pearls to pigs”? [...] Well, it would be the same as throwing pearls to pigs. I understand.

B: I’ll stick my **snout** into this conversation as well and **snatch a pearl like the snake (did) from Moses**. [...] CRUNCH, OIIIINK! Sakari (The **piggish** teleologist)

C: And again it went wrong! This snake of yours didn’t snatch any pearl from Moses, but **Moses raised a pig in the desert**. The famous **pearl belonged to Adalmina**.’

The second message contains both direct (*helmi* ‘pearl’) and indirect (*kärsä* ‘snout’, *sikaileva* ‘piggish’ and the onomatopoeic *OIIIINK!*) references to this PU. It also contains *käärme* ‘snake’ and *Mooses* ‘Moses’ – an allusion to the Old Testament story in Num. 21:4-9 about God commanding Moses to make a bronze serpent and raise it on a pole in the desert. Both allusions are blended, resulting in a situation where a snake snatches a pearl from Moses. The third message blends the same two allusions in a different way – instead of raising a snake in the desert, Moses raises a pig. This message contains yet another intertextual reference in the last sentence – *Se kuuluisa helmi kuului Adalminalle*. ‘The famous pearl belonged to Adalmina.’, which refers to the fairy tale *ADALMINAS PÄRLA* ‘Adalmina’s

<sup>209</sup> GG: sfnet.keskustelu.uskonto.kristinuskko/msg/73bd2240b829a75b, Mar 9 1999

<sup>210</sup> GG: sfnet.keskustelu.uskonto.kristinuskko/msg/f46c87271ff4ce63, Mar 9 1999

<sup>211</sup> GG: sfnet.keskustelu.uskonto.kristinuskko/msg/02ded0311c91d645, Mar 9 1999

pearl' by Zachris Topelius – a Swedish-speaking 19<sup>th</sup> century Finnish writer.

### 5.3 Variation and the PU's network structure

In the previous Section 5.2 I discussed PU cohesion, which can be regarded as the major underlying source of PU variation. In this section I will examine variation mechanisms which are licensed by the PU's network structure. Section 5.3.1 below will briefly introduce the most common morphosyntactic patterns of the *HELMIÄ SIOILLE* construction family and a possible way to describe their overlapping. In Section 5.3.2 I will look at the effect which the negative modal features appearing in the CS/PU can have on the PU's negative polarity. In 5.3.3, different types of non-default linking between syntactic constituents and phonological/conceptual structures will be examined: in 5.3.3.1 I will take a closer look at a construction family with no clear default of one particular lexical linking; Section 5.3.3.2 below will analyze non-default linking licensed by the phonological structure of a default noun in the *HELMIÄ SIOILLE* construction family; in 5.3.3.3 the examined non-defaults are licensed by LCS of default noun and verb constituents of the *HELMIÄ SIOILLE* construction family; and in 5.3.3.4 I will observe some non-default linking licensed by elements of the CS/PU (a negative modal feature and the thematic property of the argument-referent of *sika* 'pig'). In Section 5.3.4 non-default linking is explained as a result of interaction between several different construction families which share a lexical constituent and/or on the level of conceptual structure involve overlapping semantic elements.

#### 5.3.1 Morphosyntactic patterns in the *HELMIÄ SIOILLE* construction family

Table 7 below presents an overview of morphosyntactic variation in the examined construction family. It shows distribution of the most common<sup>212</sup> syntactic patterns in the *HELMIÄ SIOILLE* construction family, detected in the Google Groups corpus. A percentage number shows how frequently each pattern occurs in the entire data for this family. Morphological variation is demonstrated for the salient constituents in

---

<sup>212</sup> I.e. occurring in the data with >1% tokens.

each pattern separately. Percentage numbers given for morphological forms indicate their frequency in the respective syntactic pattern. Links between syntactic constituents and specific lexical items are not taken into account in Table 7. Thus here I am looking only at morphosyntactic patterns, which can feature both default and non-default lexical items.

Table 7 Morphosyntactic patterns of the HELMIÄ SIOILLE construction family in Google Groups data

Syntactic pattern				%	Example
<NP <sub>SUBJ</sub> >V <sub>tr</sub> NP <sub>OBJ</sub> [N <sub>1</sub> ] PP <sub>ADJUNCT</sub> [NP[N <sub>2</sub> ]]				36	
Constituent morphology	NP[N <sub>1</sub> ]	CASE	PTV	92	
			NOM	5	
			GEN	2	
			{GEN/ NOM}	1	
		NUM	PL	90	
			SG	10	
			{SG/PL}	<1	
		CASE & NUM	PTV PL	87	<i>Minä en viitsi heitellä <b>helmiä</b> sioille.</i> <sup>213</sup> lit. 'I do not bother to throw <b>pearl</b> <sub>PTV PL</sub> to pigs.'
			PTV SG	6	<i>[...] sellaista <b>helmeä</b> emme sialle heitä, emme edes (f)emäsialle.</i> <sup>214</sup> lit. '[...] we will not throw such <b>pearl</b> <sub>PTV SG</sub> to a pig, not even to a female pig.'
			NOM PL	3	<i>Ja niinhän ne siat lantalätäkköön <b>helmet</b> heittää...</i> <sup>215</sup> lit. 'And so those pigs are throwing <b>pearl</b> <sub>NOM PL</sub> into a dung pool...'
			NOM SG	2	<i>[...] mikä on se <b>helmi</b> joka on sioille jaettu tässä ryhmässä?</i> <sup>216</sup> lit. '[...] what's this <b>pearl</b> <sub>NOM SG</sub> which is distributed to pigs in this group?'
			GEN SG	2	<i>Voisitko ystävällisesti nyt heittää sen <b>helmen</b> meille sioille?</i> <sup>217</sup>

<sup>213</sup> GG: sfnet.keskustelu.evoluutio/msg/da17041bc2a6fe1b, Sep 25 2004

<sup>214</sup> GG: sfnet.keskustelu.ihmissuhteet/msg/48d031866824fb04, Jan 19 2004

<sup>215</sup> GG: sfnet.keskustelu.evoluutio/msg/3c7f1770c1ada22a, Sep 4 2004

<sup>216</sup> GG: sfnet.urheilu/msg/d49fa37eeb4a3547, Dec 11 2000



	PP[NP[N <sub>2</sub> ]]				lit. 'Could you kindly now throw this <b>pearl</b> <sub>GEN SG</sub> to us pigs?'
			{GEN SG/ NOM PL}	<1	Heitä <b>helmesi</b> päin sikojen naamaa. <sup>218</sup> lit. 'Throw <b>pearl</b> <sub>{(NOM/GEN) SG/ NOM PL}</sub> 2SG into pigs' face.'
		CASE	ALL	95	
			ILL	2	
			GEN (ETEEN)	1	
			(PÄIN) PTV	<1	
			INE	<1	
			ADE	<1	
			ABL	<1	
		NUM	PL	88	
			SG	12	
		CASE & NUM	ALL PL	85	Minä en viitsi heitellä helmiä sioille. <sup>219</sup> lit. 'I do not bother to throw pearls to <b>pig</b> <sub>ALL PL</sub> .'
			ALL SG	9	Ei, en viitsi heittää helmiä sialle minäkään. <sup>220</sup> lit. 'No, neither do I bother to throw pearls (to) <b>pig</b> <sub>ALL SG</sub> .'
			GEN SG (ETEEN)	1	Gregorius [...] jatkoi vielä letkautuksella helmistä, joita hänen ei tarvitse heitellä "saastaisen <b>sian eteen!</b> " <sup>221</sup> lit. 'Gregorius [...] continued with a quip about pearls, which he does not need to throw <b>before a filthy pig</b> <sub>GEN SG</sub> !'
			ILL PL	1	Ei sfnetin <b>sikoihin</b> helmiä haaskata, ainoastaan lokaa. <sup>222</sup> lit. 'Let's not waste pearls (on) sfnet's <b>pig</b> <sub>ILL PL</sub> , but only mud.'

<sup>217</sup> GG: sfnet.urheilu/msg/e4e0db3f16a7de45, Dec 8 2000

<sup>218</sup> GG: sfnet.huuhaa/msg/6398301113790c7c, Dec 14 1996

<sup>219</sup> GG: sfnet.keskustelu.evolutio/msg/da17041bc2a6fe1b, Sep 25 2004

<sup>220</sup> GG: sfnet.keskustelu.uskonto.kristinusko/msg/d86b88a778d780c3, Apr 9 1999

<sup>221</sup> GG: sfnet.keskustelu.uskonto/msg/c6ab0b13214142c4, Jun 27 1998

<sup>222</sup> GG: sfnet.huuhaa/msg/b938528d10680a12, Sep 13 2005

			(PÄIN) PTV SG	<1	<i>Heitä helmesi päin sikojen naamaa.</i> <sup>223</sup> lit. 'Throw your pearl(s) <b>into</b> pigs' face PTV SG.'
			ILL SG	<1	<i>Ja niinhän ne siat lantalätäkköön helmet heittää...</i> <sup>224</sup> lit. 'And so those pigs are throwing pearls (into a) <b>[puddle of dung]</b> ILL SG ..'
			INE SG	<1	<i>Sitten on ne onnettomat väliinputoajat jotka tulevat sikojen tallomiksi esitettyään helmensä väärässä paikassa.</i> <sup>225</sup> lit. 'Then there are those unlucky losers who get trampled on by pigs after having shown their pearls (in the) <b>[wrong place]</b> INE SG.'
			ABL PL	<1	<i>Varastakaa helmiä sioilta!</i> <sup>226</sup> lit. 'Steal pearls (from) <b>pig</b> ABL PL!'
			ADE PL	<1	<i>Piti vastata jotain sinulle mutta mitäpä sikoja helmillä syöttää.</i> <sup>227</sup> lit. '(I) had to give you some answer, but why feed pigs (with) <b>pearl</b> ADE PL?'
NP[N <sub>1</sub> PP[NP[N <sub>2</sub> ]]]				31	
Constituent morphology	N <sub>1</sub>	CASE	PTV	93	
			NOM	5	
			ESS	1	
		NUM	PL	96	
			SG	4	
		CASE & NUM	PTV PL	93	<i>Helmiä sioille; ei se kuitenkaan ymmärrä.</i> <sup>228</sup> lit. ' <b>Pearl</b> PTV PL to pigs; he won't understand anyway.'
			NOM SG	4	<i>Tätä uudistusta ei kuitenkaan aiota toteuttaa koska se on liian helmi sioille.</i> <sup>229</sup> lit. 'This reform is not intended to be

<sup>223</sup> GG: sfnet.huuhaa/msg/6398301113790c7c, Dec 14 1996

<sup>224</sup> GG: sfnet.keskustelu.evoluutio/msg/3c7f1770c1ada22a, Sep 4 2004

<sup>225</sup> GG: sfnet.keskustelu.uskonto.kristinuskko/msg/3b2f3302ce035c93, Oct 1 2001

<sup>226</sup> GG: sfnet.urheilu.jaakiekko/msg/c988936be180a5c4, Nov 11 1996

<sup>227</sup> GG: sfnet.tiedostot/msg/00761fcb433b406a, Oct 18 2003

<sup>228</sup> GG: sfnet.atk.sodat/msg/0030629615a7033c, Jun 14 1998

<sup>229</sup> GG: sfnet.viestinta.roskapostit/msg/77241fbcc32a559a, Jan 14 2005

					carried out because it is too much of a <b>pearl</b> <sub>NOM PL</sub> to pigs.’
			NOM PL	1	<i>Hukkaan menee kuin <b>helmet</b> sioille.</i> <sup>230</sup> lit. ‘It goes to waste like <b>pearl</b> <sub>NOM PL</sub> to pigs.’
			ESS PL	1	<i>Sinulla vain on sellainen maine, että helposti tulee tulkituksi lähes kaikki postauksesi “<b>helminä</b> sioille”, eli tylyinä ja alentuvina heittoina peeloille, jotka eivät itse viitsi ottaa asioista selvää.</i> <sup>231</sup> lit. ‘You just have such a reputation that almost all of your postings can be easily interpreted as “ <b>pearl</b> <sub>ESS PL</sub> to pigs”, i.e. rude and condescending remarks to newbies, who do not bother to find out things for themselves.’
	PP[NP[N <sub>2</sub> ]]	CASE	ALL	97	
			ILL	3	
		NUM	PL	95	
			SG	5	
		CASE & NUM	ALL PL	93	<i>Helmiä sioille; ei se kuitenkaan ymmärrä.</i> <sup>232</sup> lit. ‘Pearls (to) <b>pig</b> <sub>ALL PL</sub> ; he won’t understand anyway.’
			ALL SG	4	<i>Toki sinä et varmaan tajua logiikkaa, joten siitä puhuminen on helmiä <b>sialle</b>.</i> <sup>233</sup> lit. ‘But of course you probably don’t understand logic, so speaking of it is pearls (to a) <b>pig</b> <sub>ALL SG</sub> .’
			ILL PL	3	<i>Mustatko sitä juttua, minkä minä tuossa kevääällä lähetin tänne ryhmään otsikolla “<b>Helmii sikoihin</b>”?</i> <sup>234</sup> lit. ‘Do you remember the story which I sent to this group in the spring under the title “Pearls (into) <b>pig</b> <sub>ILL PL</sub> ”?’
			ILL SG	1	<i>Noin lupailevan tekstin postaaminen huuhaahan on kuin rekkakuormallinen</i>

<sup>230</sup> GG: sfnet.keskustelu.vitsit/msg/29310b137e70381a, Jul 3 2006

<sup>231</sup> GG: sfnet.viestinta.www/msg/bc2e61c1a2ef3bf7, Mar 28 1999

<sup>232</sup> GG: sfnet.atk.sodat/msg/0030629615a7033c, Jun 14 1998

<sup>233</sup> GG: sfnet.keskustelu.politiikka/msg/61c1aa0989a44349, Sep 4 2002

<sup>234</sup> GG: sfnet.keskustelu.uskonto.kristinusko/msg/08bc5276d8775447, Nov 10 2003

					<i>helmiä sikalan rehusiiloon.</i> <sup>235</sup> lit. ‘Posting such a promising text to huuhaa is like a truck load of pearls (into) a piggery’s [forage silo] <sub>ILL SG</sub> .’	
NEG NP <sub>OBJ</sub> [N <sub>1</sub> ] PP[NP[N <sub>2</sub> ]]				10		
	N <sub>1</sub>	CASE	PTV	100		
		NUM	PL	100		
		CASE & NUM	PTV PL	100		<i>Ei helmiä sioille, kuten sanonta kuuluu.</i> <sup>236</sup> lit. ‘No <b>pearl</b> <sub>PTV PL</sub> to pigs, as the saying goes.’
	PP[NP[N <sub>2</sub> ]]	CASE	ALL	98		
			ADE	2		
		NUM	PL	100 <sup>237</sup>		
		CASE & NUM	ALL PL	98		<i>Kyllä, ei helmiä sioille.</i> <sup>238</sup> lit. ‘Yes, no pearls (to) <b>pig</b> <sub>ALL PL</sub> ’
			ADE PL	2		<i>Ei helmiä sioilla, sanoj jo K.Vuorikin aikoinaan.</i> <sup>239</sup> lit. ‘ <b>Pig</b> <sub>ADE PL</sub> (have) no pearls, said K. Vuori once.’
NP <sub>SUBJ</sub> [N <sub>1</sub> ] V <sub>intr</sub> PP <sub>ADJUNCT</sub> [NP[N <sub>2</sub> ]]				6		
	N <sub>1</sub>	CASE	PTV	63		
			NOM	37		
		NUM	PL	83		

<sup>235</sup> GG: sfnet.huuhaa/msg/4991472cec823358, Apr 7 2003

<sup>236</sup> GG: sfnet.keskustelu.politiikka/msg/2371038599d18097, Dec 4 2001

<sup>237</sup> No singular forms of this constituent were found in my Google Groups data. However, the linking is not completely fixed, since at least two tokens of *EI HELMIÄ SIALLE* | *NEG*[*ei* 'not'] *NP<sub>OBJ</sub>*[*N<sub>PL</sub>*[*helmi* 'pearl']] *PP<sub>ALL</sub>*[*N<sub>SG</sub>*[*sika* 'pig']] | lit. 'no pearls to a pig' have been detected by performing a search in other discussion groups. E.g. in <http://kelkkalehti.com/keskustelu/index.php?topic=15838.msg602587#msg602587>, Apr 4 2010

<sup>238</sup> GG: sfnet.urheilu.jalkapallo/msg/14c6875e1dc8ac62, Nov 22 1998

<sup>239</sup> GG: sfnet.keskustelu.uskonto.kristinuskko/msg/39f6946e5e9a07ee, Aug 24 2000

PP[NP[N <sub>2</sub> ]]	CASE & NUM	SG	17	
		PTV PL	63	<i>“Liian hieno softa” IMHO, <b>helmiä</b> menee sioille.. mutta sen imutusvarmuus on kyllä parasta laatua :)</i> <sup>240</sup> lit. “Too fine an app” IMHO, <b>pearl</b> <sub>PTV PL</sub> go to pigs.. but its download reliability is really second to none :)’
		NOM PL	20	<i>Taas lensivät Tean satiirin <b>helmet</b> sioille...</i> <sup>241</sup> lit. ‘Again <b>pearl</b> <sub>NOM PL</sub> of Tea’s satire flew out to pigs.’
		NOM SG	17	<i>Nyt meni Erki <b>helmi</b> sialle.</i> <sup>242</sup> lit. ‘Now, Erki, a <b>pearl</b> <sub>NOM SG</sub> went to a pig.’
	CASE	ALL	97	
		ILL	3	
	NUM	PL	83	
		SG	17	
	CASE & NUM	ALL PL	83	<i>Tai mitä minä sitä selittämään, meni <b>helmiä</b> sioille.</i> <sup>243</sup> lit. ‘Or why should I even bother to explain it, pearls went (to) <b>pig</b> <sub>ALL PL</sub> .’
		ALL SG	13	<i>Nyt tuli <b>helmiä</b> sialle, minä ymmärrän runoutta paremmin sitä matematiikan kieltä.</i> <sup>244</sup> lit. ‘Now pearls came (to) a <b>pig</b> <sub>ALL SG</sub> , I understand poetry better than this language of mathematics.’
		ILL SG	3	<i>Menikö taas <b>helmet</b> sinne <b>kaukalo</b>on ....</i> <sup>245</sup> lit. ‘Did pearls go (to) that <b>trough</b> <sub>ILL SG</sub> again....’

<sup>240</sup> GG: sfnet.aloittelijat.kysymykset/msg/c90aae3233b1a01a, Mar 1 2001

<sup>241</sup> GG: sfnet.huuhaa/msg/53773e422c631c8a, Oct 9 1997

<sup>242</sup> GG: sfnet.keskustelu.vitsit/msg/60e595c61415385f, Sep 18 2004

<sup>243</sup> GG: sfnet.viestinta.matkapuhelimet/msg/9665e4f8b2a995b7, Feb 11 2004

<sup>244</sup> GG: sfnet.keskustelu.ihmissuhteet/msg/8d4ff6f29ec31de7, Oct 28 2003

<sup>245</sup> GG: sfnet.keskustelu.maanpuolustus/msg/27a9a4b9f4e88178, Nov 14 2003

NP[NP[N <sub>1</sub> ] DVtrN <sub>3</sub> PP[NP[N <sub>2</sub> ]]]					
			6		
	NP[N <sub>1</sub> ]	CASE	GEN	100	
		NUM	PL	100	
		CASE & NUM	GEN PL	100	
	PP[NP[N <sub>2</sub> ]]	CASE	ALL	93	
			GEN (ETEEN)	7	
		NUM	PL	83	
			SG	17	
		CASE & NUM	ALL PL	76	<i>Sinä ja muut sosialistit itse käyttäytytte niin epäkohteliaasti, epäälyllisesti ja ylimielisesti, että tarkempi analyysi olisi helmien heittämistä <b>sioille</b>.</i> <sup>246</sup> lit. 'You and other socialists yourselves behave so impolitely, non-intellectually and arrogantly that a more detailed analysis would be throwing (of) pearls (to) a <b>pig</b> <sub>ALL PL</sub> .'
			ALL SG	17	
			GEN PL (ETEEN)	7	
					<i>Käyttäydyt niin epäasiallisesti, että olisi helmien heittämistä <b>sialle</b> käyttää paljon energiaa keskusteluun.</i> <sup>247</sup> lit. 'You are behaving so inappropriately that using much energy in the debate would be throwing (of) pearls (to) a <b>pig</b> <sub>ALL SG</sub> .'
					<i>Niinpä puhuminen viestinnän todellisesta problematiikasta olisi helmien heittämistä <b>sikojen eteen</b>.</i> <sup>249</sup>

<sup>246</sup> GG: sfnet.harrastus.musiikki/msg/11d8c08c32127eb9, Sep 6 1996

<sup>247</sup> GG: sfnet.keskustelu.politiikka/msg/fb07218f428f9f34, Mar 7 2004

<sup>248</sup> GG: sfnet.keskustelu.politiikka/msg/fb07218f428f9f34, Mar 7 2004

<sup>249</sup> GG: sfnet.viestinta.www/msg/4de215cf720de09d, May 25 1997

DVtN <sub>3</sub>					lit. 'So talking about the real problems of communications would be throwing (of) pearls <b>before pig</b> <sup>GEN PL</sup> .'
		NUM	SG	100	
		CASE & NUM	PTV SG	62	<i>Inkan euroviisuehdokas oli helmien heittämistä sioille.</i> <sup>250</sup> lit. 'Inkka's Eurovision candidate was <b>throwing</b> <sup>PTV PL</sup> of pearls to pigs.'
			NOM SG	17	<i>Kuvien lähettäminen sfnettiin oli kuin helmien jakelu sioille.</i> <sup>251</sup> lit. 'Sending photos to sfnet was like <b>distribution</b> <sup>NOM SG</sup> of pearls to pigs.'
			GEN SG	7	<i>Minun Herrani on kieltänyt helmien heittämisen sikojen eteen.</i> <sup>252</sup> lit. 'My Lord has forbidden <b>throwing</b> <sup>GEN SG</sup> of pearls before pigs.'
			ELA SG	7	<i>Tähän ei sovi mikään maininta helmien heittämisestä sioille, sillä se halventaisi noita älykkäitä eläimiä.</i> <sup>253</sup> lit. 'Any mention (about) <b>throwing</b> <sup>ELA SG</sup> pearls to pigs does not fit here, because it would insult those intelligent animals.'
			ILL SG	3	<i>Vähemmän typeränä henkilönä sinulta varmaankin löytyy luonteen jaloutta viisauden helmien jakamiseen myös meille typerämmille yksilöille.</i> <sup>254</sup> lit. 'As a less stupid person you've certainly got nobility of character for <b>distributing</b> <sup>ILL SG</sup> pearls of wisdom to us more stupid individuals.'
			TRA SG	3	<i>Erkki voisi hoitaa pikimmiten Eskolalle listan yleisimmistä ilmaisuista, ellei katso sitä helmien syöttämiseksi sioille.</i> <sup>255</sup>

<sup>250</sup> GG: sfnet.harrastus.musiikki/msg/11d8c08c32127eb9, Sep 6 1996

<sup>251</sup> GG: sfnet.viestinta.nyssit/msg/0ba0179143a1926e, Sep 21 2006

<sup>252</sup> GG: sfnet.keskustelu.uskonto.kristinusko/msg/c782e1ebf6d7aa3b, Nov 3 2005

<sup>253</sup> GG: sfnet.keskustelu.varaventiili/msg/d0d357a8d8743f8a, Mar 15 2005

<sup>254</sup> GG: sfnet.keskustelu.seksi/msg/d70025cad1f8676b, Sep 8 2000

<sup>255</sup> GG: sfnet.keskustelu.vitsit/msg/1b015ffe7c5bbde6, Jul 5 2004

					lit. ‘Erkki could as soon as possible make a list of the most common expressions for Eskola, unless he considers it as <b>feeding</b> <sub>TRA SG</sub> pearls to pigs.’
<NP <sub>SUBJ</sub> >V <sub>tr</sub> NP <sub>OBJ</sub> [N <sub>1</sub> ]				3	
	NP[N <sub>1</sub> ]	CASE	PTV	83	
			GEN	8	
			NOM	8	
		NUM	PL	92	
			SG	8	
		CASE & NUM	PTV PL	83	<i>En edes viitsi heittää <b>helmiä</b>.</i> <sup>256</sup> lit. ‘I do not even bother to throw <b>pearls</b> <sub>PTV PL</sub> .’
			NOM PL	8	<i>Miten muuten ajattelit välittää nämä <b>helmet</b> internetyhteyden ja tietokoneen välityksellä?</i> <sup>257</sup> lit. ‘By the way, how did you intend to provide these <b>pearls</b> <sub>NOM PL</sub> through the Internet and computer?’
			GEN SG	8	<i>Minäpä pistän kärsäni tähänkin keskusteluun ja nappaan <b>helmen</b> niinkuin käärme Moosekselta. [...] SKRUNTS, OIIIIINK! Sakari (Sikaileva teleologisti)</i> <sup>258</sup> lit. ‘I’ll stick my snout into this conversation as well and snatch a <b>pearl</b> <sub>GEN SG</sub> like the snake (did) from Moses. [...] CRUNCH, OIIIIINK! Sakari (The piggish teleologist)’
NP[N <sub>1</sub> ] {AND/OR} NP[N <sub>2</sub> ]				1	
	N <sub>1</sub>	CASE	ELA	80	
			PTV	20	
		NUM	PL	100	

<sup>256</sup> GG: sfnet.keskustelu.politiikka/msg/2e9d4a7a89af7ec1, Apr 22 1999

<sup>257</sup> GG: sfnet.keskustelu.evolutio/msg/9684fef88f3c950e, Mar 5 2003

<sup>258</sup> GG: sfnet.keskustelu.uskonto.kristinuskko/msg/f46c87271ff4ce63, Mar 9 1999



			ELA PL	80	<i>Foneettisten merkkien käyttö historiaa käsittelevässä kirjoituksessa taitaa olla vähän yliampuva. Joku voisi jopa mutista jotain <b>helmistä</b> ja <b>sioista</b>.<sup>259</sup></i> lit. 'Using phonetic characters in a writing on history seems to be a bit extravagant. Someone could even mutter something (about) <b>pearl</b> <sup>ELA PL</sup> and <b>pig</b> <sup>ELA PL</sup> .'
		CASE & NUM	PTV PL	20	<i>Kun ei saisi olla eri mieltä kuin se 5%:n enemmistö (<b>sikoja vai helmiä</b>?) + kaiken suvaitsevaisuuden lippulaiva-helmi, Ahvenanmaa.<sup>260</sup></i> lit. 'When one should not disagree with the 5% majority (of <b>pig</b> <sup>PTV PL</sup> or <b>pearl</b> <sup>PTV PL</sup> ?) + the flagship pearl of all tolerance, Åland.'
NP[PartisP[PP[NP[N <sub>2</sub> ]] Partisr <sub>ss</sub> ] N <sub>1</sub> ]				1	
	NP[N <sub>1</sub> ]	CASE	NOM	50	
			GEN	25	
			PTV	25	
		NUM	PL	75	
			SG	25	
		CASE & NUM	GEN PL	25	<i>Ja kehoitus järjenkäyttöön olisi yhtäläinen sioille heitettyjen <b>helmien</b> kanssa.<sup>261</sup></i> lit. 'And a suggestion to use common sense would be similar to <b>pearl</b> <sup>GEN PL</sup> thrown to pigs.'
			NOM PL	25	<i>Kun vielä päälle lasketaan LenMacin muille artisteille viskomat <b>helmet</b> tyyliin Bad To Me, alkavat WTB:n coverit tuntua yhä enemmän kunnian osoitukselta/evankeliumin julistukselta/monipuolisuuden tavoittelulta.<sup>262</sup></i> lit. 'And when you add LenMac's

<sup>259</sup> GG: sfnet.keskustelu.kieli/msg/9923eb1e5a375c6b, Aug 3 2005

<sup>260</sup> GG: sfnet.keskustelu.kielipolitiikka/msg/3eea26b6cdb14cc7, Oct 17 2001

<sup>261</sup> GG: sfnet.keskustelu.vitsit/msg/59de4b1ab18f95f5, May 15 2001

<sup>262</sup> GG: sfnet.harrastus.musiikki/msg/c3f938bc588bf9df, Jan 9 1998

					<b>pearl</b> <sub>NOM PL</sub> slung to other artists in same manner as Bad To Me, WTB’s covers start to feel more and more like a tribute/gospel preaching/ pursuit of diversity.’
			NOM SG	25	<i>Possulle heitetty <b>helmi</b> tämäkin viesti, mutta kyllä sellaisia on.</i> <sup>263</sup> lit. ‘This message too is a <b>pearl</b> <sub>NOM SG</sub> thrown to a piggy, but they exist too.’
			PTV PL	25	<i>Iskarin toiminta on juurikin sellaista kuin kuvittelinkin ja pelkäänpä, että minun taidoilla lisäsäädöt olisivat kuin sioille heiteltyjä <b>helmiä</b>.</i> <sup>264</sup> lit. ‘The shock absorber’s functioning was just like I imagined and I am afraid that with my skills further adjustments would be like <b>pearl</b> <sub>PTV PL</sub> thrown to pigs.’
	PP[NP[N <sub>2</sub> ]]	CASE	ALL	100	
		NUM	PL	75	
			SG	25	
		CASE & NUM	ALL PL	75	<i>Ja kehoitus järjenkäyttöön olisi yhtäläinen <b>sioille</b> heitettyjen helmien kanssa.</i> <sup>265</sup> lit. ‘And a suggestion to use common sense would be similar to pearls thrown (to) <b>pig</b> <sub>ALL PL</sub> .’
			ALL SG	25	<i>Possulle heitetty helmi tämäkin viesti, mutta kyllä sellaisia on.</i> <sup>266</sup> lit. ‘This message too is a <b>pearl</b> <sub>NOM SG</sub> thrown to a piggy, but they exist too.’
	NP <sub>OBJ</sub> [N <sub>1</sub> ] V Partis <sub>PSS</sub> PP[NP[N <sub>2</sub> ]]				1
	NP[N <sub>1</sub> ]	CASE	NOM	50	
			PTV	50	
		NUM	PL	50	
			SG	50	

<sup>263</sup> GG: sfnet.keskustelu.seksi/msg/089f3d64e8862f08, Mar 13 2002

<sup>264</sup> GG: sfnet.harrastus.mp/msg/68945348121f4b7a, Jan 4 2005

<sup>265</sup> GG: sfnet.keskustelu.vitsit/msg/59de4b1ab18f95f5, May 15 2001

<sup>266</sup> GG: sfnet.keskustelu.seksi/msg/089f3d64e8862f08, Mar 13 2002

		CASE & NUM	NOM SG	50	<i>Ahvenanmaan kohdalla heikäläiset ovat oikeassa: <b>Helmi</b> on annettu sioille.</i> <sup>267</sup> lit. 'In Åland's case they are right: a <b>Pearl</b> <sub>NOM SG</sub> is given to pigs.'
			PTV PL	50	<i>Ja lisäksi <b>helmiä</b> ei ole tarjoiltu sioille sillä tätä [...] kampanjaa näytetään vain ja ainoastaan sivistyneille suomalaisille, ei ruotsalaiselle roskasakille joka nyt ostaa Arlaa kuitenkin.</i> <sup>268</sup> lit. 'And besides <b>pearl</b> <sub>PTV PL</sub> have not been offered to pigs because this [...] campaign is shown only to the civilized Finns, not to the Swedish riffraff that will buy Arla anyway.'
	PP[NP[N <sub>2</sub> ]]	CASE & NUM	ALL PL	100	<i>Ahvenanmaan kohdalla heikäläiset ovat oikeassa: <b>Helmi</b> on annettu <b>sioille</b>.</i> <sup>269</sup> lit. 'In Åland's case they are right: a <b>Pearl</b> <sub>NOM SG</sub> is given (to) <b>pig</b> <sub>ALL PL</sub> .'
NEG NP <sub>OBJ</sub> [N <sub>1</sub> ]				1	
	NP[N <sub>1</sub> ]	CASE & NUM	PTV PL	100	<i>Nostettaisiinko keskustelun tasoa? Alan kaipaamaan ihan oikeaa paperikirjeenvaihtoa. Ei <b>helmiä</b>....</i> <sup>270</sup> lit. 'Shall we raise the level of discussion? I am starting to miss real paper correspondence. No <b>pearl</b> <sub>PTV PL</sub> ...'

Strictly speaking, all of these patterns are regular, rule-based mappings between syntactic representations and morphological categories. There is nothing irregular, idiosyncratic in the nature of the linking itself, e.g. case marking of the noun constituents follows general rules of Finnish grammar, etc. Larger syntactic structures in which these patterns can appear are not taken into consideration in Table 7. Some instances of the verbless *NP[N<sub>1</sub> PP[NP[N<sub>2</sub>]]]* embedding into different structures have been presented in Chapter 4, Section 4.5, Examples (120) – (125), summarized in Figure 81. A relatively common structure that can feature

<sup>267</sup> GG: sfnet.keskustelu.kielipolitiikka/msg/3eea26b6cdb14cc7, Oct 17 2001

<sup>268</sup> GG: sfnet.huuhaa/msg/3723053eb53d1a86, Nov 6 1995

<sup>269</sup> GG: sfnet.keskustelu.kielipolitiikka/msg/3eea26b6cdb14cc7, Oct 17 2001

<sup>270</sup> GG: sfnet.keskustelu.skeptismi/msg/9b0f9097c3153ac4, Oct 6 2001

several of the above-presented patterns is the subject-predicate construction  $\langle NP\{W\} \rangle ON \langle \{IKÄÄN / MELKEIN / SAMA / VÄÄHÄN\} \rangle \langle NIIN \rangle KUIN Z' \langle W \rangle$  is  $\langle$  as it were / almost / same / little' like  $Z'$ , where  $Z$  is one of the following constructions:  $\langle NP_{SUBJ} \rangle Vtr NP_{OBJ} [N_1] PP_{ADJUNCT} [NP[N_2]]$  in (151),  $NP[N_1] PP[NP[N_2]]$  in (152),  $NP[PartisP[PP[NP[N_2]]] Partis_{SPSS} N_1]$  in (153) and  $NP[NP[N_1] DVtrN_3 PP[NP[N_2]]]$  in (154) below:

- (151) *Ei NT:tä ole tarkoitettu käytettäväksi telnetin kautta, joten [ssh:kin]<sub>w</sub> olisi kuin [helmiä sioille heittäisi]<sub>z</sub>.*<sup>271</sup>  
lit. 'NT is not intended to be used via telnet, so [ssh]<sub>w</sub> too would be like [throwCOND ACT SG3 pearls to pigs]<sub>z</sub>.'
- (152) *Tai sitten jotain sellaista kuin "[flarnin antaminen tuollaiselle törpölle]<sub>w</sub> on kuin [helmiä sioille]<sub>z</sub>"...*<sup>272</sup>  
lit. 'Or maybe something like "[giving flarn to a neanderthal like that]<sub>w</sub> is like [pearls to pigs]<sub>z</sub>.'
- (153) *Iskarin toiminta on juurikin sellaista kuin kuvittelinkin ja pelkäänpä, että minun taidoilla [lisäsäädöt]<sub>w</sub> olisivat kuin [sioille heitetyjä helmiä]<sub>z</sub>.*<sup>273</sup>  
lit. 'The shock absorber's functioning was just like I imagined and I am afraid that with my skills [further adjustments]<sub>w</sub> would be like [pearls thrown to pigs]<sub>z</sub>.'
- (154) *Koska käyttäjä ei ymmärrä asiasta mitään, ei häntä sovi hämmentää turhan yksityiskohtaisilla teknisillä sepustuksilla, [ne]<sub>w</sub> olisivat kuin [helmien heittämistä sioille]<sub>z</sub>.*<sup>274</sup>  
lit. 'Since the user does not understand anything about the matter, he is not to be confused by too detailed technical writings, [they]<sub>w</sub> would be like [throwing of pearls to pigs]<sub>z</sub>.'

Figure 92 below is a formal description of the first pattern in Table 7 –  $\langle NP_{SUBJ} \rangle Vtr NP_{OBJ} [N_1] PP_{ADJUNCT} [NP[N_2]]$ . Default and non-default morphology is illustrated with the help of notation developed in Chapter 3, Section 3.4.1.2.

<sup>271</sup> GG: sfnet.atk.sodat/msg/34069fb497119d86, Jun 24 1997

<sup>272</sup> GG: sfnet.viestinta.tv.babylon5/msg/ccdf2a075a971fba, Jul 15 1998

<sup>273</sup> GG: sfnet.harrastus.mp/msg/68945348121f4b7a, Jan 4 2005

<sup>274</sup> GG: sfnet.atk.nt/msg/d49cbf5a3bdbdd2a, Feb 25 1999

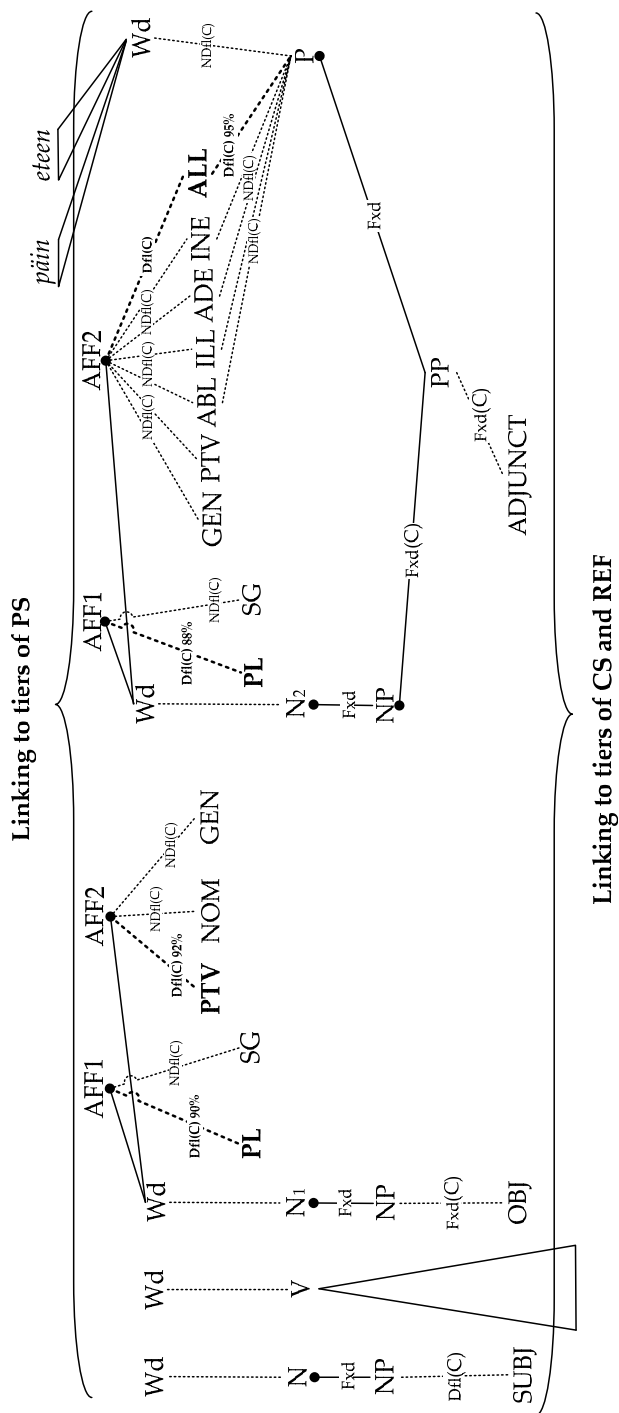


Figure 92 Morphological variation in the  $\langle NP_{SUBJ} \rangle Vtr \ NP_{OBJ} \ [N_1] \ PP_{ADJUNCT} [NP[N_2]]$  syntactic pattern (C = construction)

Patterns presented in Table 7 differ from construction-grammar metaconstructions mentioned in Section 4.5 in Chapter 4. Instead of being some abstract, schematic generalizations over substantive instances, they describe nodes and links between the tiers of syntax and morphology, which can be further connected to tiers of phonology, conceptual structure and the referential tier. The difference between the current model and the inheritance model can be demonstrated by comparing the way the transitive and the intransitive patterns are linked in Figure 93 below to Figure 80 in Chapter 4, Section 4.5, where the Transitive Construction and the Intransitive Construction are themselves represented as nodes and can only be linked via a superordinate Subject-Predicate node. In Figure 93 below one does not find any construction nodes, but instead there are network structures that overlap in some nodes (overlaps are marked with solid grey colour).

Some links to phonology had to be included in the picture – mainly to demonstrate that the same lexical item (*helmi* ‘pearl’) with the same default morphological form (*PTV PL*) is by default linked to two different syntactic positions – the subject NP constituent  $N_{1(C-II)}$  in the intransitive pattern and to the object NP constituent  $N_{1(C-II)}$  in the transitive one. Both constructions feature a PP adjunct with the same default morphology.

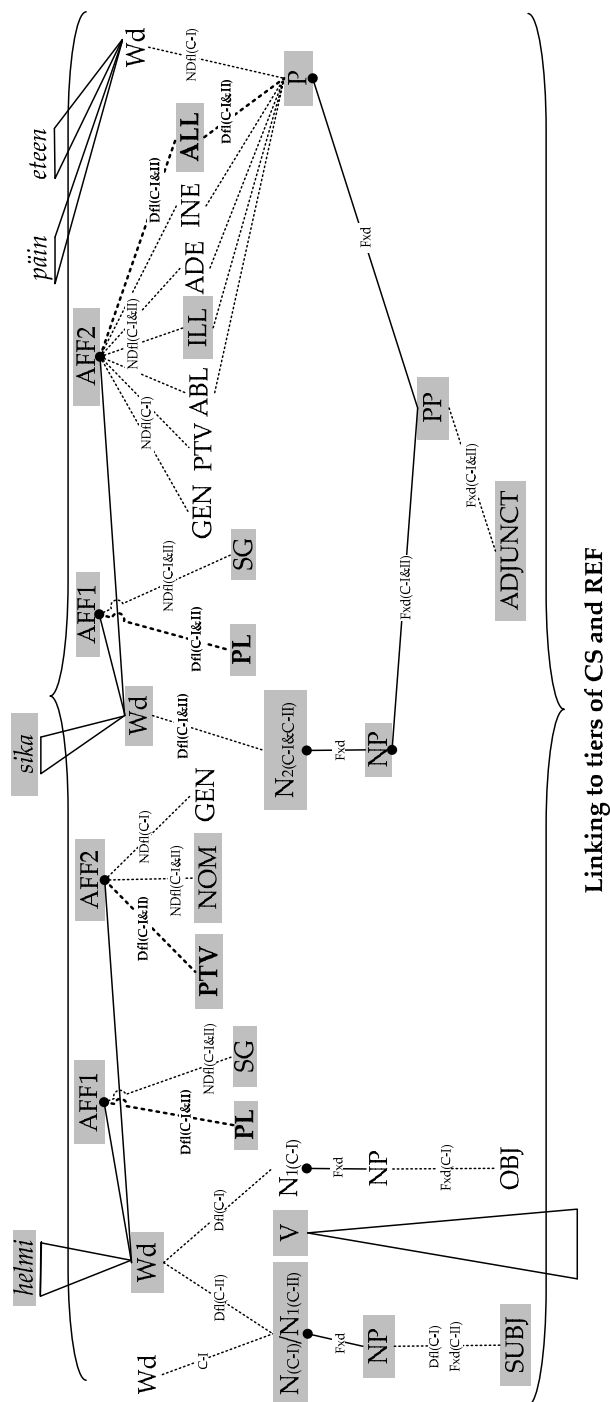


Figure 93 Overlaps between  $\langle NP_{SUBJ} \rangle Vtr NP_{OBJ} [N_1] PP_{ADJUNCT} [NP[N_2]]$  (C-I) and  $NP_{SUBJ} [N_1] Vintr PP_{ADJUNCT} [NP[N_2]]$  (C-II) patterns

### 5.3.2 The effect of the modal tier on a PU's negative polarity

In Section 3.4.3.6.2 I mentioned that the negative modality of inexpediency can be formalized by introducing the *PrNess*→*Neg* modal features where *PrNess* stands for ‘must’, ‘have to’, ‘ought to’, ‘should’ and *Neg* indicates negation. In Petrova (2009) I investigated to what extent the presence of the *PrNess*→*Neg* features in a PU's CS affects the frequency of the unit's occurrences in negative forms or negative contexts. The analysis is based on empirical data obtained from authentic postings on Usenet discussion groups by using the method described in Section 1.3 of the Introduction to this book. My hypothesis is that the presence of *PrNess*→*Neg* features in the CS of PUs with semantics of inexpediency has a direct influence on their negative polarity. The data for the *HELMIA* *SIOILLE* construction family was compared to the data for several other Finnish PUs. For the purpose of comparative analysis illustrated in Section 5.3.2.2 below, an additional database has been compiled. Ten Finnish phrasal PUs have been selected, nine of which are represented by a single construction, which is normally the default form; one is taken in its two most frequent variants – transitive and intransitive verbal constructions (items *HLPM* and *LMPM* in Table 8 below). The total of 2251 tokens are distributed between the 11 constructions as shown in Table 8 below.

Table 8 Compared constructions

PU	Abbr.	Tokens
<i>X EI TULE HULLUA HURSKAAMMAKSI Y:STÄ</i> , ‘X NEG come <sub>NEG</sub> mad <sub>PTV SG</sub> devout <sub>CMP TRA SG</sub> Y <sub>ELA</sub> ’, lit. ‘X doesn’t become more devout than a madman from Y’, id. ‘X still does not understand Y, X does not get anything out of Y’	ETHH	258
<i>X HEITTÄÄ LAPSI PESUVEDEN MUKANA</i> , ‘X throws child [bathing water] <sub>GEN SG</sub> with’, lit. ‘X throws the child out with the bathing water’, id. ‘X loses something valuable while getting rid of something unwanted’	HLPM	128
<i>X HEITTÄÄ LUSIKKA NURKKAAN</i> , ‘X throws spoon corner <sub>ILL SG</sub> ’, lit. ‘X throws the spoon into the corner’, id. ‘X dies’	HLN	102
<i>X KAIKUU KUIROILLE KORVILLE</i> , ‘X sounds deaf <sub>ALL PL</sub> ears <sub>SALL PL</sub> ’, lit. ‘X sounds for deaf ears’, id. ‘X is ignored’	KKK	163
<i>LAPSI MENEE PESUVEDEN MUKANA</i> , ‘child go <sub>SG3</sub> [bathing water] <sub>GEN SG</sub> with’, lit. ‘the child goes out with the bathing water’, id. ‘something valuable becomes lost while getting	LMPM	253



rid of something unwanted'		
<i>X MAALAA PIRUJA SEINILLE</i> , 'X paints devil <sub>PTV PL</sub> wall <sub>ALL PL</sub> ', lit. 'X paints devils on the walls', id. 'X exaggerates potential risks and possibility of failure'	MPS	282
<i>X MENEE MERTA EDEMMÄS KALAAN</i> , 'X goes sea <sub>PTV SG</sub> further fish <sub>ILL SG</sub> ', lit. 'X goes fishing further than the sea', id. 'X seeks for something farther than it is necessary'	MMEK	413
<i>X OTTAA LUSIKKA KAUNIISEEN KÄTEEN</i> , 'X takes spoon beautiful <sub>ILL SG</sub> hand <sub>ILL SG</sub> ', lit. 'X takes the spoon into the beautiful hand', id. 'X reconciles a situation'	OLKK	230
<i>X PANEE JÄITÄ HATTUUN</i> , 'X puts ice <sub>PTV PL</sub> hat <sub>ILL SG</sub> ', lit. 'X puts ice into his/her hat', id. 'X restrains one's zeal'	PJH	130
<i>X PANEE PILLIT PUSSIIN</i> , 'X puts whistle <sub>PL</sub> bag <sub>ILL SG</sub> ', lit. 'X puts whistles into the bag', id. 'X quits'	PPP	192
<i>X PITÄÄ KYNTTILÄÄ VAKAN ALLA</i> , 'X keeps candle <sub>PTV SG</sub> bushel <sub>GEN SG</sub> under', lit. 'X keeps his/her candle under the bushel', id. 'X keeps his/her talents a secret'	PKVA	100

### 5.3.2.1 Morphosyntactic means of representing the *PrNess*→*Neg* features in discourse

In this section I will concentrate on describing various ways that are used to encode linguistically the *PrNess*→*Neg* features in the Finnish language, illustrating them with examples from my database for *HHS* and some other constructions. Frequencies of occurrence of these different forms in all 12 constructions will be presented in 5.3.2.2 below and summarized in Table 11 together with other means of explicit and implicit negation.

#### 5.3.2.1.1 Modality and overt negation

On the one hand, not all means of conveying the *PrNess*→*Neg* features are explicitly negative, i.e. contain one of the overt morphosyntactic forms that are used for expressing negation in the Finnish language. On the other hand, although negation and modality often intertwine, the negative *ei* by itself does not explicitly express the *PrNess*→*Neg* features, as it can only denote that some state of affairs simply does not take place, while modality, which stands for the reason why it does not take place, remains implicit, e.g. in (155) below:

(155) *Emmehän me nyt toki helmiä sioille jakele [...]*.<sup>275</sup>

NEG<sub>PL1 hAn</sub> We now certainly pearl<sub>PTV PL</sub> pig<sub>ALL PL</sub> deal out<sub>PRES ACT NEG</sub>

<sup>275</sup> GG: sfnet.keskustelu.psychologia/msg/549093019aef9fa0, Sep 2 2001

lit. 'We are certainly **not** dealing out pearls to pigs now [...].'

The same is true for such a verb form as MA-infinitive in the abessive case (-*mAttA*) which has a negative meaning, e.g. (156) below:

- (156) *Maalailematta piruja seinille, haluaisin vain sanoa, että [...].*<sup>276</sup>

**Paint**<sub>MA-INF ABE</sub> devil<sub>PTV PL</sub> wall<sub>ALL PL</sub>, want<sub>COND ACT SG1</sub> only say<sub>A-INF</sub>, that [...].

lit. 'Without painting devils on the wall, I just want to say that [...].'

It does not express the *PrNess*→*Neg* features by itself, unless it is a part of the composite verb *olla* 'be' + *V<sub>MA-INF ABE</sub>*, which in turn is combined with a recessive modal verb (e.g. *täytyä* 'must', *pitää* 'must', *kannattaa* 'be worthwhile') or appears in a verb chain in a recessive construction, like *olla* 'be' + *ADJ PRED* + *olla* 'be' + *V<sub>MA-INF ABE</sub>*, e.g. *on pakko ~ syytä ~ hyvä ~ paras(ta) ~ tarpeellista ~ oleellista* etc. *olla nauramatta* 'it is compulsory ~ advisable ~ good ~ best ~ necessary ~ essential etc. not to laugh', e.g. in (157) and (158) below:

- (157) *Taitaa olla parasta vastaisuudessa olla maalailematta piruja seinille [...].*<sup>277</sup>

Seem<sub>PRES ACT SG3</sub> **be**<sub>A-INF</sub> **best**<sub>PTV SG</sub> future<sub>EINE SG</sub> **be**<sub>A-INF</sub> **paint**<sub>MA-INF ABE</sub> devil<sub>PTV PL</sub> wall<sub>ALL PL</sub> [...]

lit. 'It seems that in future **it is best not to paint** devils on the walls [...].'

- (158) *Oleellista on kuitenkin olla heittämättä lasta tunkiolle pesuveden mukana [...].*<sup>278</sup>

**Essential**<sub>PTV SG</sub> **be**<sub>PRES ACT SG3</sub> however **be**<sub>A-INF</sub> **throw**<sub>MA-INF ABE</sub> child<sub>PTV SG</sub> rubbish heap<sub>ALL SG</sub> bathing water<sub>GEN SG</sub> with [...].

lit. 'However **it is essential not to throw** out the child into a rubbish heap with the bathing water [...].'

### 5.3.2.1.2 Modal verbs

A highly productive way to express the *PrNess*→*Neg* features is a combination of the negative *ei* with a modal verb in negative form, e.g. (159) below:

- (159) *Ei minun kannata heittää helmiä sioille.*<sup>279</sup>

**NEG**<sub>SG3 IGEN SG</sub> **worth**<sub>PRES ACT NEG</sub> throw<sub>A-INF</sub> pearl<sub>PTV PL</sub> pig<sub>ALL PL</sub>.

lit. 'It is not worth my while to throw pearls to pigs.'

Different PUs may prefer different modal verbs. As can be seen in Table 9 below, which shows the distribution of modal verbs in the *HHS* data, the

<sup>276</sup> GG: sfnet.harrastus.lemmikit.koirat/msg/f73db2add93bd1d4, Dec 3 2001

<sup>277</sup> GG: sfnet.harrastus.autot/msg/cc6c3972442a409a, Dec 8 2000

<sup>278</sup> GG: sfnet.keskustelu.varaventiili/msg/2eda26af9b689a3d, Oct 11 2001

<sup>279</sup> GG: sfnet.keskustelu.filosofia/msg/bdf29b56c9c77e80, Apr 22 2006

verb *kannattaa* ‘be worth’ is the most frequent for this PU. However, it is not the case for MMEK, which has *tarvita* ‘need’ as the most used modal verb, as shown in Table 10 below.

Table 9 Modal verbs in the HHS data

MV in HHS	Tokens	%
<i>kannattaa</i> ‘be worthwhile’	9	36
<i>viitsiä</i> ‘be bothered’	8	32
<i>saada</i> ‘be allowed to’	2	8
<i>suostua</i> ‘agree’	2	8
<i>kuulua</i> ‘have to, behoove’	1	4
<i>tarvita</i> ‘need’	1	4
<i>haluta</i> ‘want’	1	4
<i>aikoa</i> ‘intend’	1	4

Table 10 Modal verbs in the MMEK data

MV in HHS	Tokens	%
<i>tarvita</i> ‘need’	76	52
<i>kannattaa</i> ‘be worthwhile’	51	35
<i>pitää</i> ‘must’	8	6
<i>viitsiä</i> ‘be bothered’	5	3
<i>haluta</i> ‘want’	2	1
<i>aikoa</i> ‘intend’	1	1
<i>meinata</i> ‘mean’	1	1
<i>taitaa</i> ‘seem’	1	1

### 5.3.2.1.3 Negative imperative forms

The negative *ei* has imperative forms for SG2 (*älä*) and PL2 (*älkää*) while the finite verb appears in the negative form like in (160) below:

- (160) *Koska he eivät keskustele täällä keskustelualueella, niin **älä** sinäkään enää jaa helmiä sioille täällä.*<sup>280</sup>

Because they NEG<sub>PL3</sub> converse this<sub>SG</sub> discussion area<sub>SG</sub>, so NEG<sub>IMPV</sub> ACT SG2 you<sub>KAA</sub>n anymore distribute<sub>PRES</sub> ACT NEG pearl<sub>PTV</sub> PL pig<sub>ALL</sub> PL here.

lit. ‘Because they don’t converse in this discussion area, so **don’t** you either distribute pearls to pigs here anymore.’

<sup>280</sup> GG: finet.harrastus.rautatiet/msg/607f6209f92d5556, Jan 26 1999

Negative imperative for PL1 is expressed in colloquial Finnish by the negative *ei* plus a finite verb in the passive negative form, e.g. (161) below:

- (161) *Toisaalta, ei tämän yhteiskunnan ihmisillä ole paljon aivoja rakentamaan mitään parempaa, joten ei<sup>281</sup> heitetä turhaan helmiä sioille.*

On the other hand, NEG this<sub>GEN SG</sub> society<sub>GEN SG</sub> people<sub>ADE PL</sub> be<sub>PRES ACT NEG</sub> much brain<sub>PTV PL</sub> build<sub>MA-INF ILL</sub> anything<sub>PTV SG</sub> good<sub>CMP PTV SG</sub>, so NEG<sub>PA</sub> throw<sub>PRES PSS NEG</sub> in vain pearl<sub>PTV PL</sub> pig<sub>ALL PL</sub>.

lit. 'On the other hand, people of this society don't have much brain to build anything better, so **let's not** throw pearls to pigs in vain.'

The imperative mood represents deontic modality. In Hakulinen et al. (2004: 1511) it is said that directive or permission expressed by the imperative is always originated by the situation and the utterer, while deonticity indicated by other means has its source in norm, custom, law or some other authority outside the situation. This does not seem to apply to PUs. It is true that the choice of the PU by the speaker and the use of the imperative mode are motivated by the situation, like in (160) and (161) above. However it would be inaccurate to say that deontic modality in this case comes only from the situation, as I assume here that it is inherent to the PU.

#### 5.3.2.1.4 Necessive construction *EI OLE X{SYYY, TARPEEN, AIKA, TARKOITUS, PYRKIMYS ETC} V<sub>INF</sub>* 'there is no X{reason, necessary, time, intention, striving} to ...'

According to Hakulinen et al. (2004: 1502) this construction too expresses deontic modality, e.g. in (162) below:

- (162) *Mutta ei ole "tarpeen" maalata pirua seinälle, [...].*<sup>282</sup>

But NEG be<sub>PRES ACT NEG</sub> necessary paint<sub>A-INF</sub> devil<sub>PTV SG</sub> wall<sub>ALL SG</sub>, [...].

lit. 'But it is not necessary to paint a devil on the wall, [...].'

<sup>281</sup> GG: sfnet.keskustelu.huumeet/msg/8b4d54ae19a71366, May 20 1998

<sup>282</sup> GG: sfnet.keskustelu.uskonto.kristinuskko/msg/a4445760db503aeb, May 15 2006

### 5.3.2.1.5 Implicitly negative forms

#### 5.3.2.1.5.1 Rhetorical questions

Rhetorical questions are utterances that are formally interrogative, but do not have an information-seeking function and are not meant to be answered. Instead they rather imply statements. According to Kiuru (1977: 60), rhetorical questions in positive form usually correspond to negative sentences and vice versa, e.g. (163) – (165) below:

(163) *Miksi antaa helmiä sioille.*<sup>283</sup>

Why give<sub>EA-INF</sub> pearl<sub>PTV PL</sub> pig<sub>ALL PL</sub>.

lit.: Why give pearls to pigs.

(164) *Vai mitä sitä helmiä sioille heittelemään, vai mitä?*<sup>284</sup>

Or what<sub>PTV</sub> this<sub>PTV SG</sub> pearl<sub>PTV PL</sub> pig<sub>ALL PL</sub> throw<sub>MA-INF ILL</sub>, or what<sub>PTV</sub>?

lit. 'Or why throw pearls to pigs, or what?'

(165) [...] *mutta pitääkö tässä nyt sitten lähteä merta edemmäs kalaan...*<sup>285</sup>

[...] but must<sub>PRES ACT KO</sub> this<sub>SINE SG</sub> now then go<sub>A-INF</sub> sea<sub>PTV SG</sub> further fish<sub>ILL SG...</sub>?

lit. '[...] but now then, must one go fishing further than the sea...?'

#### 5.3.2.1.5.2 Implicitly negative verbs, e.g. *varoa* 'beware of', *välttää* 'avoid' etc.

(166) *Varo vaan heittelemästä helmiä sioille!*<sup>286</sup>

Beware<sub>IMPV ACT SG2</sub> just throw<sub>MA-INF ELA</sub> pearl<sub>PTV PL</sub> pig<sub>ALL PL</sub>!

lit. 'Just beware of throwing pearls to pigs!'

#### 5.3.2.1.5.3 Construction (ON) *TURHA(A) V<sub>INF</sub>* 'it is unnecessary, useless, pointless to ...'

This implicitly negative construction is a negative counterpart for the necessary modal verb *kannattaa* 'be worthwhile' (Hakulinen et al. 2004: 1482):

(167) *On kuitenkin turha heittää helmiä sioille.*<sup>287</sup>

Be<sub>PRES ACT SG3</sub> however pointless throw<sub>A-INF</sub> pearl<sub>PTV PL</sub> pig<sub>ALL PL</sub>.

lit. 'However, it is pointless to throw pearls to pigs.'

(168) *Turhaa heittää helmiä sioille :-)*<sup>288</sup>

<sup>283</sup> GG: sfnet.keskustelu.seksi/msg/517691b01058d2da, Feb 7 1997

<sup>284</sup> GG: sfnet.keskustelu.laki/msg/1e076c9181cc9afc, Jun 8 1998

<sup>285</sup> GG: sfnet.harrastus.ruoka+juoma/msg/33b6a2aea9546e0c, Jan 21 2003

<sup>286</sup> GG: sfnet.keskustelu.seksi/msg/ff7d795e7d22edfe, Nov 17 1997

<sup>287</sup> GG: sfnet.viestinta.nyssit/msg/b4f5d187a8a445cc, Aug 28 2005

**Pointless**<sub>PTV SG</sub> **throw**<sub>A-INF</sub> **pearl**<sub>PTV PL</sub> **pig**<sub>ALL PL</sub>.  
lit. 'It is pointless to throw pearls to pigs.'

#### 5.3.2.1.5.4 Construction *SUOTTA V<sub>INF</sub>* 'in vain, unnecessarily to ...'

(169) *Suotta heittää lasta pesuveden mukana....*<sup>289</sup>

**In vain throw**<sub>A-INF</sub> **child**<sub>PTV SG</sub> **bathing water**<sub>GEN SG</sub> **with...**

lit. 'It is in vain to throw the child out with the bathing water...'

#### 5.3.2.2 Frequencies of negation

According to Hakulinen et al. (1980: 120) 9% of clauses in a Finnish running text (neutral documentary prose, 10149 clauses, and 66,851 words) contain negation expressed by *ei*. Insofar as my data represents another text genre, I performed a separate quantitative analysis of negation in a sample consisting of 29 whole Usenet messages (700 graphic sentences, 9569 words). According to my estimates about 30% of the sentences contain various types of negation. The question, which I have formulated, was whether or not the presence of the *PrNess*→*Neg* features in a PU's semantic structure will influence the frequency of its occurrences in negative form/negative contexts. To answer this question the data containing *HSS* was compared to the 11 constructions, presented in Table 8 above.

The *HSS* construction family was represented in this analysis by its transitive construction <*NP<sub>SUBJ</sub>*>*Vtr NP<sub>OBJ</sub> [N<sub>1</sub>] PP<sub>ADJUNCT</sub>[NP[N<sub>2</sub>]]*. It was essential to take into account that different constructions of the *HSS* construction family exhibit different degrees of ability to undergo negation. The choice of the 11 other items can be explained as follows. *ETHH* can be regarded as a negative polarity item due to the fact that its form contains explicit negation. It was chosen in order to find out whether there is any difference in negation frequency between a genuine negative polarity PU and PUs with the *PrNess*→*Neg* features. *HLPM*, *LMPM*, *MPS*, *MMEK* and *PKVA* were chosen as conceivably similar to *HHS*, i.e. they do not have explicit negation in form, but include implicit *PrNess*→*Neg* features in their semantic structure. Together with the *HHS* they constitute the first control group (G1), and according to my hypothesis will show quite similar results. The second control group (G2) includes *HLN*, *KKK*, *OLKK*, *PJH* and *PPP*, which contain neither

<sup>288</sup> GG: sfnet.atk.laitteet.pc/msg/4b2dc68f14567706, Feb 13 2001

<sup>289</sup> GG: sfnet.keskustelu.politiikka/msg/bf7e4b75d91e8133, Feb 4 2002

negation nor  $PrNess \rightarrow Neg$  modal features and were chosen to demonstrate the influence of the  $PrNess \rightarrow Neg$  features on negation frequency.

Different types of explicit and implicit negation, which were counted, are shown in Table 11 below. Both the construction-internal and external scope of negation was considered. For each construction two values have been obtained representing:

- Percentage of all negative tokens in the item's database.
- Percentage of item-internal negation in all negative tokens.

These pairs of values were compared with the help of a scatter chart presented in Figure 94 below. The chart shows that there is a clear gap between G1 and G2. The average difference is 66% on the scale of construction-internal negation (value X axis) and 42% on the scale of both internal and external negation (value Y axis). There is also some difference between *ETHH* and G1: on average it is 16% for the value X axis and 28% for the value Y axis.

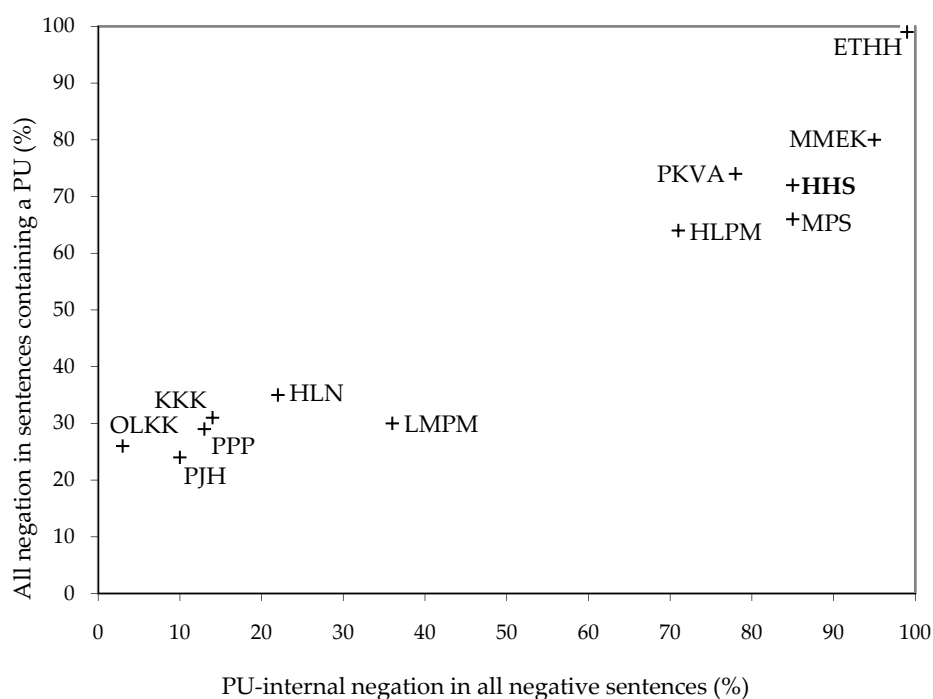


Figure 94 Negation frequencies in compared constructions

Besides this evidence for obvious influence of the  $PrNess \rightarrow Neg$  features on negation frequency I found another interesting result: despite this

modality, *LMPM* positioned itself among the items of G2. This fact points to the conclusion that the conditions in which a certain construction is used play an important role in its ability to undergo negation. My data shows that *LMPM* is used primarily as a statement or prediction of an event: 'X happened/happens/will happen'. The items in G1 and G2 also differ in regard to what types of construction-internal negation can be applied to them. Table 11 below presents a summarized overview of frequencies of different types of explicit and implicit negation that have been detected in my data.

Table 11 Different types of explicit and implicit negation summarized

PU	Types of construction-internal negation (%)										
	ABE	EI	ei ole X	EI + MV	IE	NEG IMPV	NEG IMPV PSS	INEGV	RQ	turha	suotta
ETHH	1	97							2		
<b>G1</b>											
MMEK	1	3	1	46		1	4		32	11	1
MPS	5	12	3	37		17	18		3	5	
HHS	1	26		35		12	4	1	16	5	
PKVA		43		10		24		2	21		
HLPM	2	26	3	36		7	24				2
<b>G2</b>											
LMPM		85		4	7			4			
HLN		62		25	13						
KKK		100									
PPP		57	14	29							
PJH		100									
OLKK		50	50								

*ABE* = *MA* infinitive in the abessive case, *MV* = modal verb, *IE* = *ILMAN ETTÄ* 'without', *NEG IMPV* = negative imperative forms of *EI*, *NEG IMPV PSS* = negative imperative expressed by passive, *INEGV* = implicitly negative verb, *RQ* = rhetorical question

According to Kurikka (1979), in 76% of cases negation is expressed in Finnish by *ei*. The average frequency of *ei* negation among the items in G2 is indeed 76%. This was the only type available for two of them (*KKK* and *PJH*). In G1 this average was considerably lower, only 22%, as it yields the palm to those forms that express the *PrNess*→*Neg* features. G2 items, on



the contrary, use very few of these forms. Thus, the results of the quantitative analysis of negation in the examined 12 constructions show a considerable difference that exists between PUs with the *PrNess*→*Neg* features and those that do not have them in their semantic structure. Apparently, there seems to be a clear connection between the presence of the *PrNess*→*Neg* features and the degree of the PU's negative polarity, although conditions in which a certain construction is used represent another important factor.

### 5.3.3 Default and non-default linking between a PU's syntactic constituents, phonology and LCS (lexical variation)

Following Jackendoff, I do not assume that lexical items are inserted into syntactic derivations; instead they establish a tripartite correspondence between a syntactic constituent, a specific phonological form and a specific conceptual structure. In this section I will focus on such correspondence in the PU's construction family. In Figure 95 below default links between syntactic constituents, phonological structures and lexical conceptual structures are emphasized. Consequently, the same mapping, but with a non-default value, will represent the phenomenon, which in the mainstream phraseology is often referred to as lexical variation (or lexical substitution), i.e. non-default lexical items appearing in the same syntactic positions as the default ones.

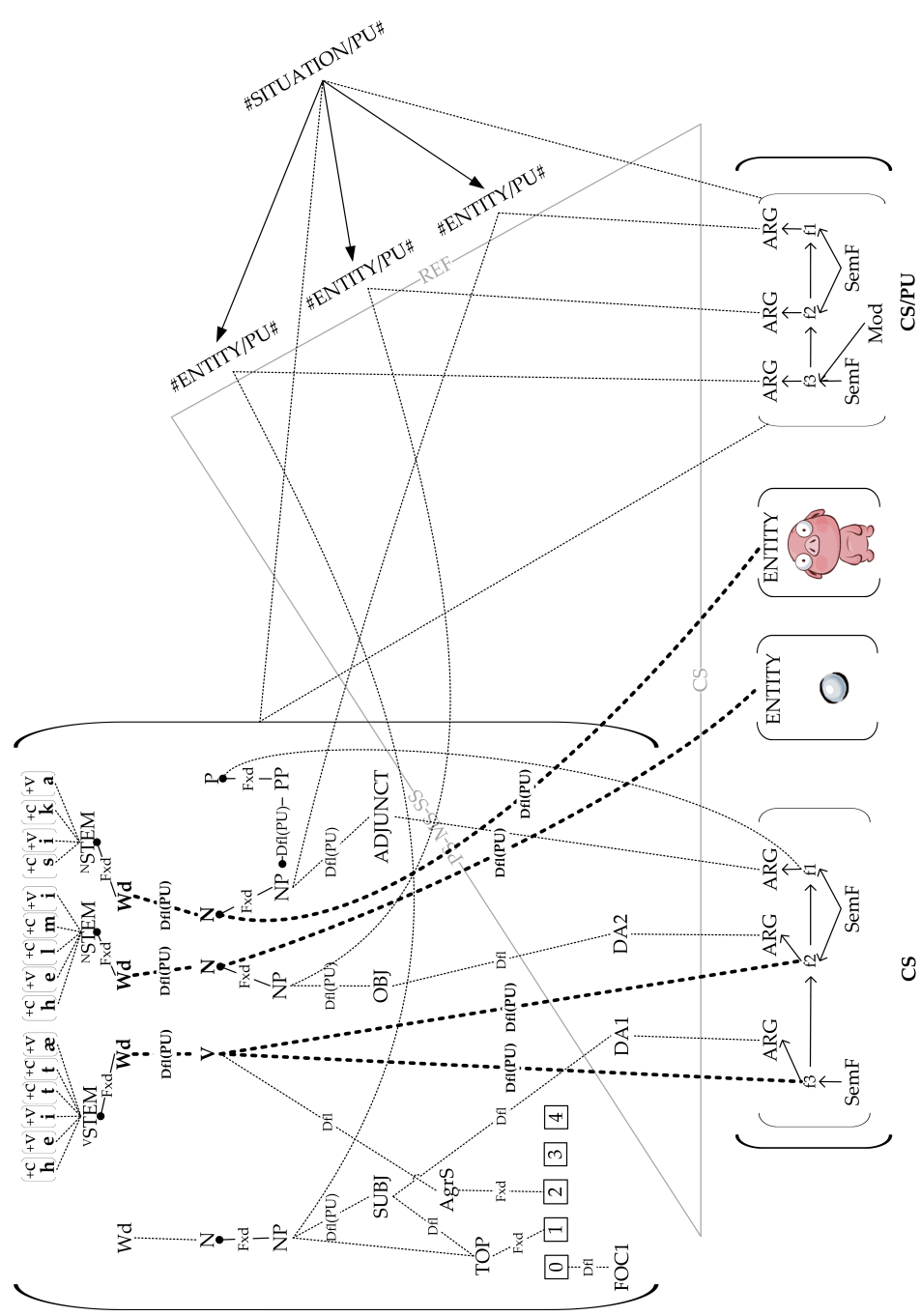


Figure 95 Correspondence linking between syntactic constituents, phonology and conceptual structure

In phraseological theory, lexical variation in PUs has been for a long time regarded as a purely occasional phenomenon, peculiar to an author's individual style. This can be explained by the fact that before large computer-based corpora have become available the overwhelming majority of research into PU usage was based on data collected from fictional literature. Observations of lexical variation made by Mokienko (1980: 25) on the basis of dialectological material demonstrated the universality of this phenomenon. Yet, even in more recent corpus-based studies there is a tendency to label a large portion of non-default lexical items as idiomatic wordplay as opposed to systematic idiom variation as in, e.g., Langlotz (2006), reviewed earlier in Chapter 4, Section 4.2.2.3 of this book. I find such a distinction rather obscure and misleading, since even wordplay is systematic in the sense that it still follows the rules set by the language system as a whole, i.e. it conforms to general principles of syntactic-phonological-conceptual interface.

In what follows, I will look at default and non-default correspondences between a PU's syntactic constituents and phonological structures in the construction families of two Finnish PUs:

- *X HEITTÄÄ HELMIÄ SIOILLE* | *NP<sub>SUBJ</sub>[N{X}] V{heittää 'throw'} NP<sub>OBJ</sub>[N<sub>PL</sub>{helmi 'pearl'}] PP<sub>ALL</sub>[N<sub>PL</sub>{sika 'pig'}]* | lit. 'X throws pearls to pigs', id. 'X causes a transfer of some entity, evaluated by the speaker as good, to a recipient, evaluated as inadequate in some way' (henceforth abbreviated as HHS);
- *X <EI> YMMÄRRÄ/TIEDÄ Y:STÄ/Y:N PÄÄLLE YHTÄ PALJON/ENEMPÄÄ KUIN SIKA HOPEALUSIKASTA* | *NP<sub>SUBJ</sub>[N{X}] <NEG{ei 'not'}> Vcogn{ymmärtää 'understand'/tietää 'know'} PP<sub>ELA/PÄÄLLE</sub>[NP<sub>PTV</sub>{Y} COMP{yhtä paljon kuin 'as much as'/enempää kuin 'more than'} NP<sub>SUBJ</sub>[N{sika 'pig'}] PP<sub>ELA</sub>[NP{hopealusikka 'silver spoon'}]* | lit. 'X <does not> understand(s) Y /know(s) about Y as much as/more than a pig about a silver spoon', id. 'X does not understand Y/know about Y at all' (henceforth abbreviated as KSH).

Figure 96 below shows distribution of relative frequencies for default and non-default lexical items in these construction families: from the *HHS* construction family there are correspondences of both noun constituents *N{helmi 'pearl'}* and *N{sika 'pig'}* and the verb *V{heittää 'throw'}*, and from the latter correspondences of the second noun constituent *N{hopealusikka 'silver spoon'}*<sup>290</sup>. Relative frequencies have been calculated based on the

<sup>290</sup> For now, I will just follow Kari (1993) and assume that *hopealusikka* 'silver spoon' is the default lexical item for this syntactic position. However, later in Section 5.3.3.1 of this chapter this assumption will be tested against a corpus of 117 tokens gathered from Google Groups. In this data, the most frequent item is *satelliitti* 'satellite'.

number of occurrences which every constituent has in the entire data for the respective construction family, excluding cases where a construction does not include this particular constituent. The entire data for the *HHS* family includes 480 tokens; however, the number of each constituent's actual occurrences is different: 271 tokens for *V{heittää 'throw'}*..., 476 tokens for *N{helmi 'pearl'}*... and 457 tokens for *N{sika 'pig'}*.... Thus, the relative frequency 0.34 for the *V{heittää 'throw'}* stands for 93 default tokens out of 271; 0.94 for *N{helmi 'pearl'}* for 446 default tokens out of 476 and 0.92 for *N{sika 'pig'}* for 420 default tokens out of 457.

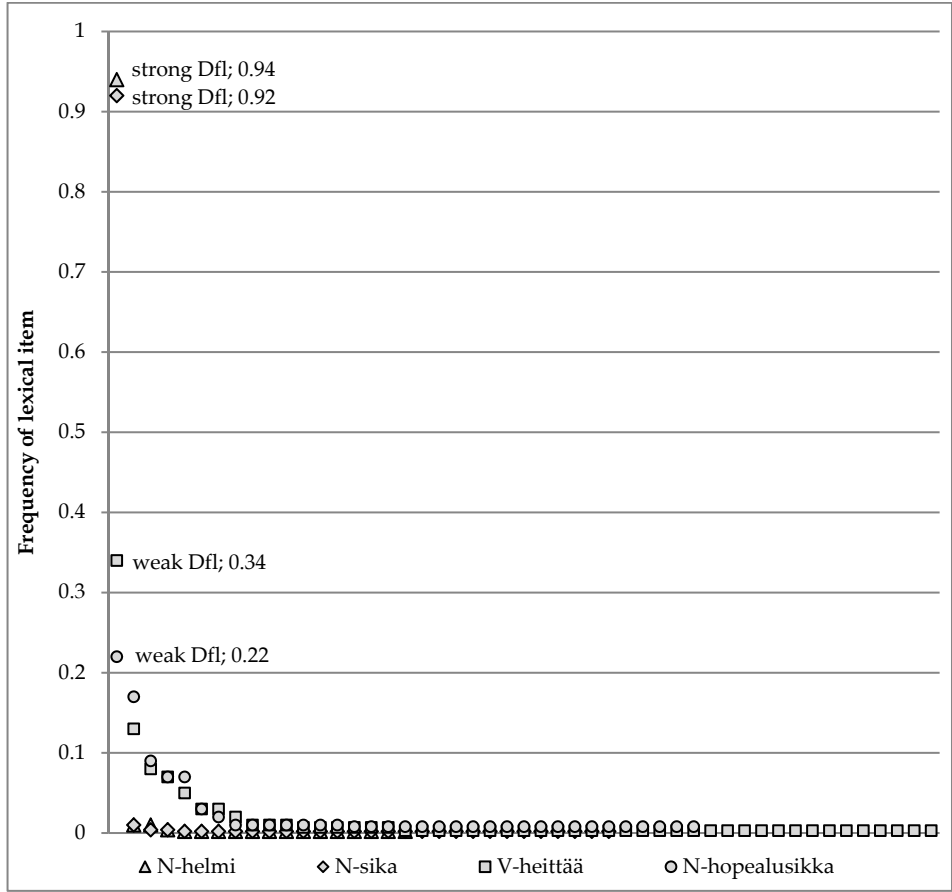


Figure 96 Distribution patterns for default (Dfl) and non-default lexical items in *HHS* and *KSH* construction families

Table 12 Table 12 below presents correlations between these patterns. The strongest correlation (0.999973) can be detected between frequencies of lexical correspondences of the *HHS* constituents *N{helmi 'pearl'}* and

*N{sika 'pig'}*, while correspondences of the *HHS* constituent *V{heittää 'throw'}* strongly correlate in their frequencies with those of the *KSH*-constituent *N{hopealusikka 'silver spoon'}* (0.946943). The weakest correlation is observed for correspondences of the *KSH* constituent *N{hopealusikka 'silver spoon'}* and those of the *HHS* constituents *N{helmi 'pearl'}* and *N{sika 'pig'}* (0.724412 and 0.720208901 respectively).

Table 12 Correlations in frequencies

	HHS <i>N{helmi 'pearl'}</i>	HHS <i>N{sika 'pig'}</i>	HHS <i>V{heittää 'throw'}</i>	KSH <i>N{hopealusikka 'silver spoon'}</i>
HHS <i>N{helmi 'pearl'}</i>	1			
HHS <i>N{sika 'pig'}</i>	0.999973	1		
HHS <i>V{heittää 'throw'}</i>	0.908601	0.896755592	1	
KSH <i>N{hopealusikka 'silver spoon'}</i>	0.724412	0.720208901	0.946943	1

As can be seen in Figure 96 above, an average 0.65 difference exists between the default connections' values of the two correlating pairs: the *HHS* constituents *N{helmi 'pearl'}* and *N{sika 'pig'}* both have an average default link strength of 0.93, while the average default strength of the *HHS* constituent *V{heittää 'throw'}* and the *KSH* constituent *N{hopealusikka 'silver spoon'}* is 0.28. Another striking discrepancy can be observed in the distances between the most frequent correspondence and the one following it. For the *HHS* constituents *N{helmi 'pearl'}* and *N{sika 'pig'}* the average distance is 0.92 while for the *HHS* constituent *V{heittää 'throw'}* and the *KSH* constituent *N{hopealusikka 'silver spoon'}* it is 0.13. It seems that we are dealing with two different default patterns – the strong default and the weak one. It is also interesting to compare frequencies of default vs. non-default lexical realizations of syntactic constituents in *HHS* (Figure 96 above) to those of each constituent's occurrence vs. absence in the entire data for this construction family (N=480), including constructions whose syntactic structure does not feature one (or two) of the constituents in question (Figure 97 below). Again, the *HHS* constituents *N{helmi 'pearl'}* and *N{sika 'pig'}* with their respective frequencies of 98% and 94% are in the clear lead, while *V{heittää 'throw'}* is the least frequent constituent in the family – it appears in only 55% out of a total of 480 *HHS* tokens.

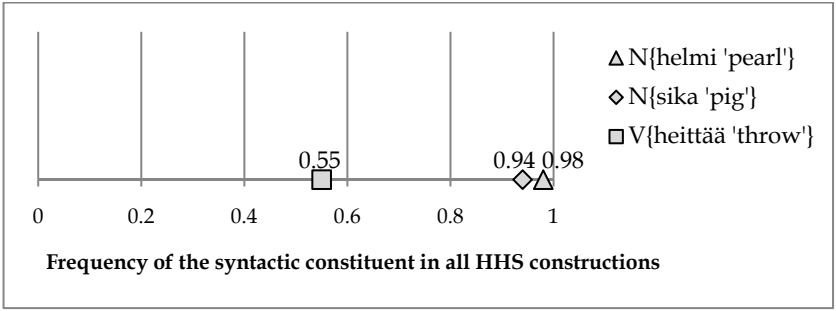


Figure 97 Relative frequencies of syntactic constituents in the HHS construction family

One could also combine both frequencies on horizontal and vertical axes, as in Figure 98 below.

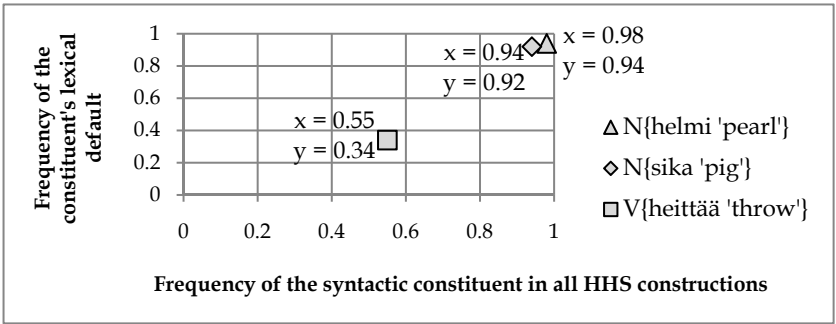


Figure 98 Combined frequencies of syntactic constituents and their lexical defaults in the HHS construction family

In Petrova (2007a) the same criteria that are combined in the above picture were analyzed separately in order to detect the key constituent of HHS. The difference between *V{heittää 'throw'}* and both noun constituents is quite apparent. However, the question of whether the difference between *N{helmi 'pearl'}* and *N{sika 'pig'}* is statistically significant remains open. Also the scale of linking strength presented in Chapter 3, Section 3.4.2, Figure 54 needs more precise, statistically based criteria of default linking value compared to free, unrestricted mapping, strong vs. weak default patterns, etc. No conclusive statements on these matters can be made at this point, since the hypothesis needs to be tested with the help of statistical methods and against a larger data set including a control group of non-idiomatic regular expressions. This task falls outside the scope of the current study and thus must be left for future research. Statistically based definition of categories and boundaries is a

research topic of its own. A solid methodological framework is offered by e.g. Arppe (2008), who describes a whole range of statistical methods for studying linguistic alternations with multiple outcomes.

Figure 99 below demonstrates the default linking for nouns in the *HHS* construction family. The number which follows the default value label indicates the strength of this particular linking.

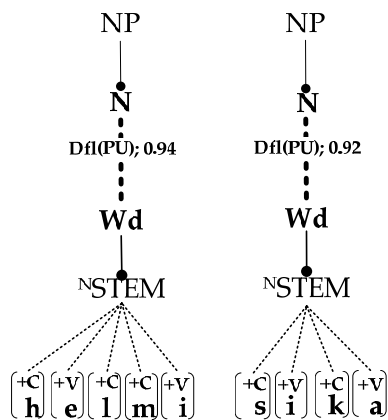


Figure 99 Default linking for noun constituents in the *HHS* construction family

Figure 100 below attempts to summarize the share of different types of non-default lexical linking for these two syntactic constituents.

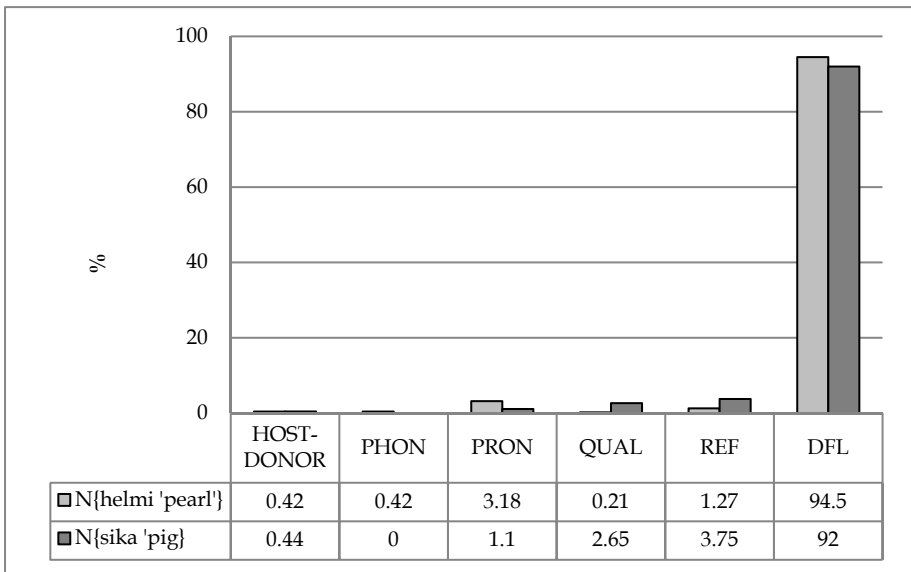


Figure 100 Default (DFL) and non-default realizations of noun constituents in the *HHS* construction family.

Underlying mechanisms for non-default linking between a PU's syntactic constituents and phonological structures will be discussed in Sections 5.3.3.2 – 5.3.3.4, Section 5.3.4 and Section 5.4. In Section 5.3.3.2 it is presented as being motivated by the phonological structure of the default lexical item (the type marked as PHON in Figure 100 above). In 5.3.3.3 non-default lexical linking to phonology is analyzed as being licensed by the default item's LCS – qualia structure (QUAL) of nouns (5.3.3.3.1) and the argument structure of the verb (5.3.3.3.2). In 5.3.3.4 non-default correspondence of the verb to phonology is assumed to be licensed by the conceptual structure of the phraseological unit (CS/PU). Section 5.3.4 examines non-default lexical linking licensed by interaction (HOST-DONOR) between PUs of different construction families and Section 5.4.2 looks at referentially licensed (REF) linking of syntactic constituents.

As long as I am mostly interested in semantic aspects underlying non-default lexical correspondence, linking to a relative pronoun in embedding (Moon 1998: 110), i.e. relegation of a part of an idiom to a relative clause dependent on a noun constituent, will not be considered here, since it does not bring any new semantic elements to the picture. Formally, the modified noun does not appear with other constituents within the structure of the same clause. Instead it is replaced by a relative pronoun, referring back to the antecedent of the relative clause like in Examples (170) and (171) below. Examples like these have to be treated separately, because, unlike non-default correspondences that lead to certain semantic changes in the construction, relative pronouns have no independent meaning. The same holds for demonstrative pronouns that anaphorically refer to a lexical constituent, which appears in the preceding clause like in Example (172) below:

- (170) [...] mikä on se **helmi joka** on sioille jaettu tässä ryhmässä?<sup>291</sup>  
lit. '[...] what's this **pearl which** is distributed to pigs in this group?'
- (171) [...] eri mieltä kanssasi olevien kirjoittajien leimaaminen [...] epäsuorasti **sioiksi joille** jaat helmiä [...] <sup>292</sup>  
lit. '[...] indirectly branding writers that do not share your opinion as **pigs to whom** you distribute pearls [...]'
- (172) Onhan meillä laatuakin...vai luuletko että 6-päinen peliporukka muuten kokoontuisi 2-3 kertaa viikossa neljän seinän sisälle...tosin mahtuuhan sekaan

<sup>291</sup> GG: sfnet.urheilu/msg/0a3b6adb727ae07d, Dec 11 2000

<sup>292</sup> GG: sfnet.keskustelu.uskonto.kristinusko/msg/383433edbf826257, Oct 12 1999



*paljon paskaakin, mutta löytyy seasta helmiäkin...eikä niitä ole heitetty säkeissä ostetuille sioille.*<sup>293</sup>

lit. 'After all, we have also quality ... or do you think that a 6-member group of gamers would otherwise meet 2-3 times a week inside four walls ... although indeed a lot of shit fits in as well, but **pearls** can be found too ... and **these** have not been thrown to pigs bought in pokes'.

Non-defaults of this type are marked as PRON in Figure 100 above. On the other hand, linking to personal and other pronouns, which can be regarded as referring to an entity in the #SITUATION#/PU rather than simply to the lexical constituent, like in Example (173) below, has been included in the category of referentially licensed non-default lexical linking marked as REF in Figure 100 above and analyzed in Section 5.4.2 in connection to discourse-related variation:

(173) *Vai onko tämä jotain henkilökohtaista vittuilua, kun heittelet tietämyksen helmiä minulle?*<sup>294</sup>

lit. 'Or is this some kind of personal aggravation, throwing pearls of knowledge to **me**?'

### 5.3.3.1 *Hopealusikka* 'silver spoon' or *satelliitti* 'satellite'? An example of weak default lexical linking

So far, I have discussed non-default correspondences of syntactic constituents to phonological and conceptual structures in the *HHS* construction family. In this section, I will examine another construction family, which exhibits a considerable degree of flexibility in its lexical correspondences. In Section 5.3.4 below, for the sake of simplicity, correspondence between the noun constituent in the second PP<sub>ELA</sub>, phonological form *hopealusikka* and lexical conceptual structure [SILVER SPOON] is treated as the default mapping in two constructions of the same construction family – the negative *X EI YMMÄRRÄ/TIETÄÄ Y:STÄ ENEMPÄÄ KUIN SIKA HOPEALUSIKASTA* | NP<sub>SUBJ</sub>[N{X}] NEG{ei} Vcogn{ymmärtää 'understand', tietää 'know'} PP<sub>ELA</sub>/PÄÄLLE/NP<sub>PTV</sub>{Y} COMP{enempää kuin 'more than'} NP<sub>SUBJ</sub>{sika 'pig'} PP<sub>ELA</sub>[NP{hopealusikka 'silver spoon'}] | lit. 'X does not understand Y/ know about Y more than a pig about a silver spoon', and the affirmative *X YMMÄRTÄÄ/TIETÄÄ Y:STÄ YHTÄ PALJON KUIN SIKA HOPEALUSIKASTA* | NP<sub>SUBJ</sub>[N{X}] Vcogn{ymmärtää 'understand'/ tietää 'know'} PP<sub>ELA</sub>/PÄÄLLE/NP<sub>PTV</sub>{Y} COMP{yhtä paljon kuin 'as

<sup>293</sup> GG: sfnet.harrastus.pelit.rooli/msg/1ae150ef94865b10, Jan 8 1999

<sup>294</sup> GG: sfnet.keskustelu.uskonto.kristinusko/msg/18c570fc6e57bb53, Apr 25 2005

*much as*'} NP<sub>SUBJ</sub>{*sika* 'pig'} PP<sub>ELA</sub>[NP{*hopealusikka* 'silver spoon'}] | lit. 'X understands Y/ knows about Y as much as a pig about a silver spoon' both meaning 'X does not understand Y/ know about Y at all'. However, a closer look at empirical data for this construction shows that *hopealusikka* and [SILVER SPOON] are in fact linked to a syntactic position which allows linking to multiple phonological and conceptual structures. The question of which of these links is a default correspondence for this syntactic constituent is not a trivial one.

Consider the fact that different dictionaries present different variants of the same construction: the dictionary of modern Finnish idioms contains the entry *EI YMMÄRRÄ ENEMPÄÄ KUIN SIKÄ HOPEALUSIKASTA* lit. 'does not understand more than a pig **about a silver spoon**' (Kari 1993); the collection of Finnish sayings from the Nivala region (Takalo and Junttila 1979) lists the form *EI YMMÄRRÄ ENEMPÄÄ, KU SIKÄ TUULIMYLLYSTÄ* lit. 'does not understand more than a pig **about a windmill**'; and the Book of Finnish Folk Similes (Kuusi 1960) presents it as *EI TIEDÄ ENEMPÄÄ KUIN SIKÄ POHJANTÄHDESTÄ* lit. 'does not know more than a pig **about the North Star**'.

A search in a digital collection of Finnish folk sayings, dating from the 1930s (Sananparsikokoelma 2010) returned the following variants presented in (174) – (182) below:

- (174) *EI TIEDÄ ENEMPÄÄ KUIN SIKÄ SUNNUNTAISTA*  
lit. 'does not know more than a pig **about Sunday**' (1 token);
- (175) *EI YMMÄRRÄ ENEMPÄÄ KUIN SIKÄ POHJATUULESTA*  
lit. 'does not understand more than a pig **about the north wind**' (1 token);
- (176) *EI TIEDÄ ENEMPÄÄ KUIN SIKÄ SATULASTA*  
lit. 'does not know more than a pig **about a saddle**' (1 token);
- (177) *EI TIEDÄ ENEMPÄÄ KUIN SIKÄ SODASTA*  
lit. 'does not know more than a pig **about war**' (1 token);
- (178) *EI {YMMÄRRÄ/TIEDÄ} ENEMPÄÄ KUIN SIKÄ {PÄIVÄNNOUSUSTA/ PÄIVÄNNOUSUA}*  
lit. 'does not {understand/ know} more than a pig **about the sunrise**' (3 tokens);
- (179) *EI {YMMÄRRÄ/TIEDÄ} ENEMPÄÄ KUIN SIKÄ {PYHÄPÄIVÄSTÄ/ PYHÄSTÄ PÄIVÄSTÄ/ PYHÄSTÄ}*  
lit. 'does not {understand/ know} more than a pig **about a holy day**' (5 tokens);
- (180) *EI {YMMÄRRÄ/TIEDÄ} ENEMPÄÄ KUIN SIKÄ POHJANTÄHDESTÄ*

lit. 'does not {understand/ know} more than a pig **about the North Star**' (2 tokens);

- (181) *EI {YMMÄRRÄ/TIEDÄ/ÄLYÄ} ENEMPÄÄ KUIN SIKA HOPEALUSIKASTA*

lit. 'does not {understand/ know} more than a pig **about a silver spoon**' (18 tokens);

- (182) *EI {YMMÄRRÄ/TIEDÄ} ENEMPÄÄ KUIN SIKA {TUULIMYLLYSTÄ/ TUULIMYLLYN PÄÄLLE}*

lit. 'does not {understand/ know} more than a pig **about a windmill**' (22 tokens, including the affirmative construction).

Empirical data collected from the World Wide Web and Usenet discussion groups shows that both constructions – the negative *X EI YMMÄRRÄ/TIETÄÄ Y:STÄ ENEMPÄÄ KUIN SIKA HOPEALUSIKASTA* | *NP<sub>SUBJ</sub>[N{X}] NEG{ei} Vcogn{ymmärtää 'understand', tietää 'know'} PP<sub>ELA</sub>/PÄÄLLE/NP<sub>PTV</sub>{Y} COMP{enempää kuin 'more than'} NP<sub>SUBJ</sub>{sika 'pig'} PP<sub>ELA</sub>[NP{hopealusikka 'silver spoon'}]* | lit. 'X does not understand/know about Y more than a pig about a silver spoon' and the affirmative *X YMMÄRTÄÄ/TIETÄÄ Y:STÄ YHTÄ PALJON KUIN SIKA HOPEALUSIKASTA* | *NP<sub>SUBJ</sub>[N{X}] Vcogn{ymmärtää 'understand'/ tietää 'know'} PP<sub>ELA</sub>/PÄÄLLE/NP<sub>PTV</sub>{Y} COMP{yhtä paljon kuin 'as much as'} NP<sub>SUBJ</sub>{sika 'pig'} PP<sub>ELA</sub>[NP{hopealusikka 'silver spoon'}]* | lit. 'X understands/knows about Y as much as a pig about a silver spoon' – show a great deal of lexical flexibility almost in every syntactic position, but especially in the linking to the second PP<sub>ELA</sub> (Table 13 below).

Table 13 Lexical variation in the KSH construction family

The target of comparison			Relation		The basis of comparison			
NP <sub>SUBJ</sub>	NEG	V <sub>cogn</sub>	PP	<AdvP>	COMP	<AP>	NP <sub>SUBJ</sub>	PP
X	Neg-C	ei 'no'	Neg-C	<yhtään sen> enempää 'none the more'				elektroniikka 'electronics', golfmailat 'golf clubs', hammasharja 'toothbrush', hopea 'silver', hopealautanen 'silver plate', hopealusikka 'silver spoon', hopeamitali 'silver medal', hulipuna 'lipstick', höyrypannu 'steam pan', Joulu 'Christmas', jänis 'hare', jääkiekko 'ice hockey', kello 'clock, watch', kuu 'moon', kvantifysiikka 'quantum physics', lentäminen 'flying', mikroaaltouuni 'microwave oven', mopedi 'moped', Nintendo, pankkikortti, 'bank card', pasta 'pasta', pianonsoitto 'piano playing', poljentahti 'polestar', politiikka 'politics', polkupyörä 'bicycle', poltto-moottori 'combustion engine', pontikka 'moonshine', positiini 'porcelain', pottuhalme 'land reclaimed by burning-over and placed under spud' pyhä 'holiday', sarjakuvat 'comics', satelliitti 'satellite', Shakespeare, samppanja ynnä mansikat 'champagne and strawberries', sinfonia 'symphony', stereot 'stereo (player)', sähkö 'electricity', taivaantähdet 'heaven's stars', taivas 'sky, heaven', tekokuu 'satellite', tinalusikka 'pewter spoon', tuulimylly 'windmill', tähtitiede 'astronomy', täytekakku 'layer cake', videot 'video (player)', ydinfysiikka 'nuclear physics' ydinsukellusvene 'nuclear submarine'
				edes {sen<kään> vertaa/sitä} 'even that much'				
	Affirm-C		Affirm-C	vain vähän enemmän 'just a little more'	kuin 'as, like, than'	keskiverto 'average'	apina 'monkey'	
				vähemmän 'less' {saman/sen} verran 'equally' yhtä {paljon/ vähän} 'as much/little'	mitä 'what'	proverbiaalinen 'proverbial' se kuuluisa 'this famous'	hullu 'loony' kala 'fish' lehmä 'cow' pääsi 'ram' sika 'pig'	

The lexical items appearing in the PP<sub>ELA</sub> adjunct position in the basis of comparison range between various domains: e.g. luxuries (*hopealautanen* 'silver plate', *posliini* 'porcelain'), technology (*kello* 'clock', *sähkö* 'electricity', *mikroaaltouuni* 'microwave oven'), food (*täytekakku* 'cake', *shampanja ynnä mansikat* 'champagne and strawberries'), celestial body (*kuu* 'Moon', *tähti* 'star'), science (*ydinfysiikka* 'nuclear physics', *tähtitiede* 'astronomy'), culture and sports (*jääkiekko* 'ice-hockey', *sinfonia* 'symphony') etc. These categories are of course very approximate. Adjuncts in the target and the basis of comparison can belong to the same domain (in Example (183) below to alcoholic drinks and in Example (184) to culture), or different domains (in Example (185) to technology and culture and in Example (186) to technology and food):

- (183) *Eihän tuo tietysti onnistu silloin, kun vanhemmatkaan eivät ymmärrä väkijuomista enempää kuin sika pontikasta.*<sup>295</sup>  
lit. 'That will of course fail if parents do not understand **about the spirits** any more than a pig **about moonshine**.'
- (184) *Ja joo, jos sä oot niin läpitönkkö juntti ettet sä tajuu jostain oopperasta ja baletista enempää ku sika Shakespearesta, niin ei se merkitse sitä, että kaikki suomalaiset (etenkään stadilaiset) olis samanlaisia pölkkyjä kuin sä oot.*<sup>296</sup>  
lit. 'And yeah, if you're such a total stiff hillbilly that you don't get **things like opera and ballet** any more than a pig gets **Shakespeare**, it does not mean that all Finns (especially Helsinki natives) would be the same blockheads as you're.'
- (185) *Visual C++ virheenkorjaus näyttää otsikon ilmoituksen. Sen enempää en siitä irti saa, kun ymmärrän ohjelmasta saman kuin sika sinfoniasta.*<sup>297</sup>  
lit. 'Visual C++ debugging displays a header declaration. I can't get any more out of it than that, since I understand about **the program** as much as a pig about **a symphony**.'
- (186) *Kyllä kai minä ymmärrän, että sinä tajuat prosessiteollisuudesta suunnilleen yhtä paljon kuin sika täytekekusta, joten et tietenkään tarkoittanut heitollasi mitään muuta kuin "(ydin)sähkömiehet hölmöjä, minä viisas".*<sup>298</sup>  
lit. 'I think I understand that you know about the **processing industry** about as much as a pig about **a layer cake**, so of course you did not mean anything else with your remark than "(nuclear) electricians – fools, me – wise.'"

<sup>295</sup> GG: sfnet.keskustelu.yhteiskunta/msg/6b04e9a25a61c36a, Oct 1 1998

<sup>296</sup> GG: sfnet.keskustelu.asuminen/msg/451817fec3d16a7, Feb 10 2004

<sup>297</sup> GG: sfnet.atk.ohjelmointi/msg/840726bd631969a1, Jul 29 2002

<sup>298</sup> GG: sfnet.atk.sodat/msg/ba991732994deafe, Feb 5 2001

The list of lexical variants within the construction family presented in Table 13 above and Figure 101 below does not pretend to be exhaustive. Figure 101 illustrates a distribution of 14 lexical correspondences of the NP constituent in the PP<sub>ELA</sub> adjunct in the basis of comparison, which occur with >1 tokens each (96 tokens altogether) among the total of 117 relevant tokens of the construction family in question, manually gathered out of a total of 473 hits returned by advanced searches in Google Groups from *sfnet*. messages posted between 1996 and 2005 and containing the exact wording “*kuin sika*” ‘like a pig’:

1. *satelliitti* ‘satellite’ (26 tokens; 22.2%)
2. *hopealusikka* ‘silver spoon’ (20 tokens; 17%)
3. *jääkiekko* ‘ice hockey’ (11 tokens; 9.4%)
4. *tuulimylly* ‘windmill’ (9 tokens; 7.6%)
5. *hopealautanen* ‘silver plate’ (9 tokens; 7.6%)
6. *pohjantähti* ‘North Star’ (4 tokens; 3.4%)
7. *helmi* ‘pearl’ (3 tokens; 2.5%)
8. *höyrypannu* ‘steam boiler’ (2 tokens; 1.7%)
9. *jänis* ‘hare’ (2 tokens; 1.7%)
10. *<digitaali>kello* ‘<digital> watch’ (2 tokens; 1.7%)
11. *pottuhalme* ‘land reclaimed by burning-over and placed under spud’ (2 tokens; 1.7%)
12. *taivaan tähti* ‘heaven’s star’ (2 tokens; 1.7%)
13. *taivas* ‘heaven, sky’ (2 tokens; 1.7%)
14. *video* ‘video’ (2 tokens; 1.7%)



This data indicates that the lexical variant  $PP_{(ELA/PÄÄLLE)}\{\textit{satelliitti} ‘satellite’} might actually be the default PP for this construction in modern Finnish (or at least in the context of computer-mediated communication). This is a rather curious fact, since the dictionary of modern Finnish idioms (Kari 1993) lists this construction in the form *EI YMMÄRRÄ ENEMPÄÄ KUIN SIKA HOPEALUSIKASTA* ‘does not understand more than a pig about a silver spoon’. It is also interesting to compare these results with an older (dating from the 1930s) dialectal material from *Sananparsikokoelma* (2010), where the  $PP_{(ELA/PÄÄLLE)}\{\textit{tuulimylly} ‘windmill’} occurs with a larger number of tokens than the lexical variant  $PP_{(ELA/PÄÄLLE)}\{\textit{hopealusikka} ‘silver spoon’}.$$$

### 5.3.3.2 Non-default lexical linking licensed by a default constituent’s phonological structure

The type of lexical variation marked in Figure 102 as PHON represents tokens in Examples (187) and (188) below. Here, in (187) the constituent *helmi* ‘pearl’ is substituted by the capitonym Finnish female proper name *Helmi* and in (188) by *hedelmä* ‘fruit’, which in PL PTV form bears phonological resemblance to the noun *helmi* ‘pearl’ in the same form:

(187) *Väkivaltaa sarjakuissa: Heikki heitti **Helmiä**<sup>299</sup> sioille.*<sup>300</sup>

lit. ‘Violence in comic books: Heikki threw **Helmi** to pigs.’

<sup>299</sup> *Heikki* and *Helmi* are fictional characters appearing in *HELMI JA HEIKKI* ‘Helmi and Heikki’, which is a Finnish translation of the American comic strip *BLONDIE*.

<sup>300</sup> GG: sfnet.keskustelu.kieli.kaantaminen/msg/75d751e67ea2db4d, Jan 15 2001



NPsubj{X} V{heittää 'throw'} NP<sub>1</sub>OBJPTV[N<sub>PL</sub>{*helmi* 'pearl'}] PP<sub>ALL</sub>[NP<sub>PL</sub>{*sika* 'pig'}]

NPsubj{Heikki} V<sub>PAST ACT SG3</sub>{heittää 'throw'} NP<sub>2</sub>OBJPTV[N<sub>SG</sub>{*Helmi* 'Pearl'}] PP<sub>ALL</sub>[NP<sub>PL</sub>{*sika* 'pig'}]

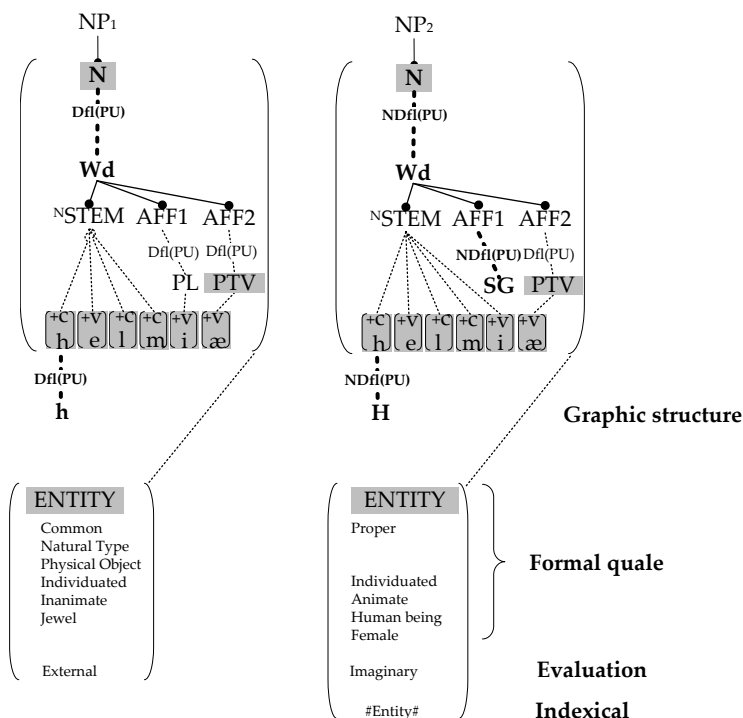


Figure 102 Formal description of non-default noun linking in (187) compared to the default one

(188) Välillä jo luulin että tyhjyyteen kirjoitellaan. Kivaa sekin olisi ollut, ei olisi ainakaan jaeltu *hedelmiä* sioille.<sup>301</sup>

lit. 'At times I thought that we are writing into emptiness. That would have been nice too, at least **fruit** would not have been distributed to pigs.'

<sup>301</sup> GG: sfnet.huuhaa/msg/1abb81200bfe2e64, Aug 1 2001

NPsubj[X] V{heittä 'cast'} NP<sub>1</sub>OBJ PTV[NrL{helmi 'pearl'}] PP<sub>ALL</sub>[NP<sub>PL</sub>{sika 'pig'}]  
NEG{ei} COP V<sub>COND</sub>ACT NEG{olla 'be'} ADV {ainakaan 'at least'} V<sub>PAST</sub>PSS NEG{jaella 'distribute'} NP<sub>2</sub>  
OBJ PTV[NrL{hedelmä 'fruit'}] PP<sub>ALL</sub>[NP<sub>PL</sub>{sika 'pig'}]

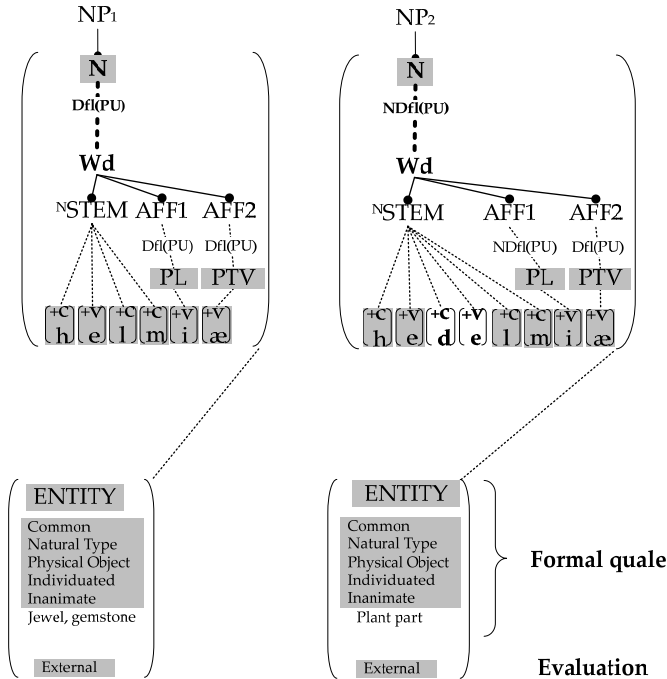


Figure 103 description of non-default noun linking in (188) compared to the default one

Examples (187) and (188) above demonstrate a slightly different degree of similarity between the default and non-default items. In Example (187) (Figure 102), both phonological structures are identical, with some differences in morphology and quite distinct conceptual structures. In Example (188) (Figure 103), morphological structures are identical, conceptual structures feature some common elements in the formal quale and valuation, while the segmental structure of the non-default item includes two phonemes ('d' and 'e') more than that of the default one. Note that Figure 102 and Figure 103 do not contain the complete semantic descriptions of the lexical items in question, but only the elements that are necessary to demonstrate overlaps and differences between their structures.

### 5.3.3.3 Non-default lexical linking licensed by a default constituent's LCS

#### 5.3.3.3.1 Non-defaults licensed by a default noun's LCS

In this section I will discuss non-default linking of syntactic constituents in the *HHS* construction family to phonological structures of lexical items, which are semantically related to the default lexical item *sika* 'pig'. Semantic relations (including synonymy, hyponymy, hypernymy, metonymy, holonymy etc.) that exist between the default and non-default items will be analyzed in accord with the theory of compositional semantics of nouns proposed by Pustejovsky (1995, 2001, 2003), namely qualia structure, which is described in more detail in Chapter 3, Section 3.4.1.3.1 of this book. In this case, non-default correspondence is assumed to be licensed by the links that can be established between the elements of the qualia structures of the default lexical item and those of its non-default substitutes.

Examples (189) – (199) below demonstrate some non-default lexical realizations of the  $PP_{ADJUNCT}[NP[N_2]]$  in the *HHS* construction family:

- (189) *Hänen kohdallaan noudatan sitä vanhaa neuvoa, jonka mukaan helmiä ei pidä antaa eräille **eläimille**, jotka eivät niistä ymmärrä.*<sup>302</sup>  
lit. 'In his case I follow that old advice according to which pearls should not be given to certain **animals** that do not understand about them.'
- (190) *Vaikka helmien heittelyistä kaikenlaisille **kotieläimille** on veistetty sananlaskuja, niin kyllä tässä mm. kotieläinen tarkkailijat helmien arvoja laskea osaavat.*<sup>303</sup>  
lit. 'Although proverbs have been thought up about tossing pearls to all kinds of **domestic animals**, here domestic animal-watchers among others are able to calculate the value of pearls.'
- (191) *... mutta jos voit heittää mielestäsi joitakin helmiä, niin nakkaappa tälle **karjulle** jokunen.*<sup>304</sup>  
lit. '... but if you think you can throw some pearls, then chuck a few to this **boar**.'
- (192) *Vai heittelenkö taas turhaan helmiä **porsaille**?*<sup>305</sup>  
lit. 'Or am I again throwing pearls to **piglets** in vain?'
- (193) ***Possulle** heitetty helmi tämäkin viesti, mutta kyllä sellaisia on. Aivan varmasti.*<sup>306</sup>

<sup>302</sup> GG: sfnet.keskustelu.psykologia/msg/84a3a7b166fee04b, Feb 2 1999

<sup>303</sup> GG: sfnet.keskustelu.evolutio/msg/af7c2720ed58e544, Jul 23 2003

<sup>304</sup> GG: sfnet.harrastus.kulttuuri.sarjakuvat/msg/fe50541c78c3f5ed, May 11 1998

<sup>305</sup> GG: sfnet.keskustelu.ihmissuhteet/msg/ac84d8165094ba8f, Apr 9 2000

lit. 'This message too is a pearl thrown to a **piggy**, but they exist too. For sure.'

- (194) *Onneksi UJT:n helmistä piisaa myös joulupossujen käyttelyyn myös.*<sup>307</sup>  
lit. 'Fortunately, there are enough of UJT's pearls for the needs of **Christmas piggies**, too.'
- (195) *Menikö taas helmet sinne kaukaloon....*<sup>308</sup>  
lit. 'Did pearls go into that **trough** again....'
- (196) *A: Miten nuo helmet liittyy tähän viestiin?*  
*B: Taisivat lipsahtaa possujen kaukaloon :-)*<sup>309</sup>  
lit. 'A: How are these pearls related to this message?  
B: I think they slipped into the **piggies' trough** :-)'
- (197) *Noin lupailevan tekstin postaaminen huuhaahan on kuin rekkakuormallinen helmiä sikalan rehusiiloon.*<sup>310</sup>  
lit. 'Posting of such a promising text to huuhaa is like a truck load of pearls into a **piggery's forage silo**.'
- (198) *Kun vastassa on ymmärtämättömyys, kokemuksen puute ja haluttomuus tarkastella asioita avoimesti ja ennakkoluulottomasti, perustelun 'helmiä' ei näille saparohännille kannata kädestään näyttää.*<sup>311</sup>  
lit. 'When faced with lack of understanding, lack of experience and unwillingness to look at things openly and open-mindedly, it's pointless to show 'pearls' of explanation to these **short tails**.'
- (199) *Ja niinhän ne siat lantalätäkköön helmet heittää...*<sup>312</sup>  
lit. 'And so the pigs are throwing pearls **into a puddle of manure**...'

Figure 104 below shows a semantic network created by constituency links between the qualia structure of the default noun *sika* 'pig'(a) appearing in the *PP<sub>ADJUNCT</sub>[NP[N<sub>2</sub>]]* position and qualia structures of its substitutes presented in (189) – (199) above. Each quale is formalized as a value of a variable, while the latter represents the entity itself. Variable indices (a)-(k) are assigned as follows:

1. Formal quale (F): *sika* 'pig'(a) is a kind of *kotieläin* 'domestic animal'(b) (190), which in turn is a kind of *eläin* 'animal'(c) (189). *Karju* 'boar'(j) (191) is a kind of *sika* 'pig'(a).

<sup>306</sup> GG: sfnet.keskustelu.seksi/msg/089f3d64e8862f08, Mar 13 2002

<sup>307</sup> GG: sfnet.keskustelu.energia/msg/c764e7864b5592ee, Dec 22 2005

<sup>308</sup> GG: sfnet.keskustelu.maanpuolustus/msg/27a9a4b9f4e88178, Nov 14 2003

<sup>309</sup> GG: sfnet.tori.myydaan.atk.komponentit/msg/b3880fa624e6109d, Feb 23 2003

<sup>310</sup> GG: sfnet.huuhaa/msg/4991472cec823358, Apr 7 2003

<sup>311</sup> GG: sfnet.keskustelu.psykologia/msg/2b2aa217882b69bc, Mar 28 1999

<sup>312</sup> GG: sfnet.keskustelu.evolutio/msg/3c7f1770c1ada22a, Sep 4 2004

2. Agentive quale (A): *porsas* 'piglet'(e) (192) and *possu* 'piggy'(e) (193) are (to\_be(a)), i.e. will grow into a *sika* 'pig'(a). *Lanta* 'dung, manure'(k) (199) denotes animal(c) feces and comes into existence as a product of animals' vital functions.
3. Telic quale (T): a pig is kept for its meat and is usually slaughtered before Christmas and traditionally served at the Christmas table as *joulupossu* 'Christmas piglet'(f) (i.e. ham) (194); *kaukalo* 'trough, manger'(g) (195), (196) is a vessel used for feeding/watering domestic animals; *sikala* 'piggery'(h) (197) functions as a dwelling place for pigs; *rehusiilo* 'forage silo'(i) (197) is used for storing fodder.
4. Constitutive quale (C): *saparohäntä* lit. 'a short tail'(d) is a part of a pig's body, but the word is also metonymically used to denote a pig (198).

The relationship between the nodes in Figure 104 below is assumed to be based on constituency (represented by a solid line with a round dot pointing at the constituent node). Thus, the notation  $C(f) \bullet - C(a)$  indicates that the Constitutive quale of the variable (f) (representing the entity *joulupossu* 'Christmas piglet') is a part of the Constitutive quale of the variable (a) (representing the entity *sika* 'pig'). As one can see, the overlapping elements do not necessarily belong to the same quale, cf.  $T(g) \bullet - F(a)$  where  $F(a)$  is an element in the taxonomic structure of the word *pig* encoding that it is a kind of cattle and  $T(g)$  in the Telic quale of the *trough* gives information about its characteristic function – being a drinking vessel for cattle.

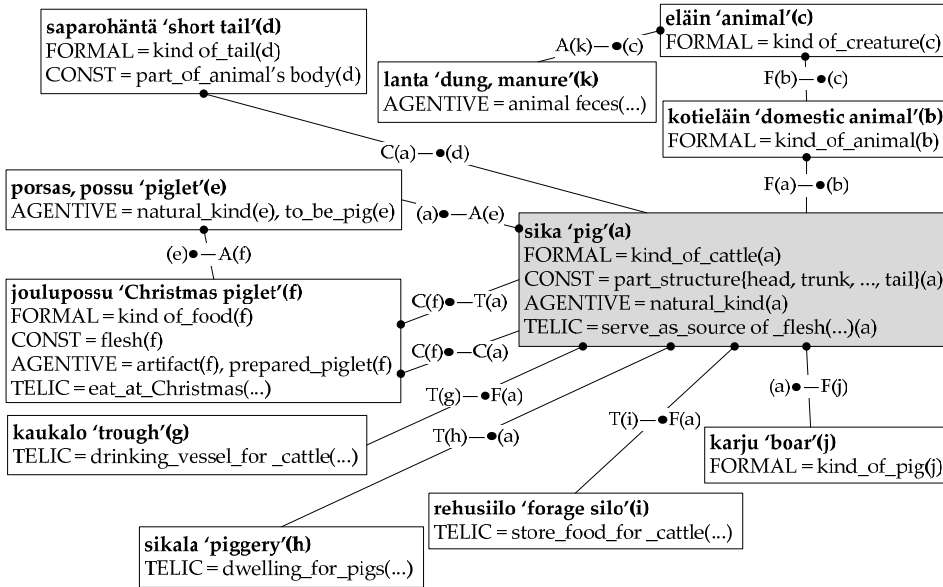


Figure 104 A possible semantic network based on the qualia structure of the lexical item *SIKA* 'pig'

The semantic model presented in Figure 104 above is rather tentative and includes only some of the possible links; its further development was not possible within the limits of the current study. The central idea, however, seems to be clear: the semantic analysis based on qualia structure allows accounting in a systematic way for a whole range of semantic relations, which can exist between default and non-default lexical items.

Non-default lexical elements licensed by the qualia structure of the item *sika* 'pig'(a) can appear in other syntactic positions as well. In (200) – (206) below one can see tokens of the PU with a non-default linking of the verb constituent to the phonological/conceptual structures of the lexical item *syöttää* 'feed', which can be related to *sika* 'pig'(a) via the Telic quale of the thematic argument FOOD, which appears in the LCS of this verb (Figure 105 below):



lit. 'It turned out to be quite long. Read and enjoy, I do not write to these very often. Pearls to pigs – when you **feed** pearls to a pig, it is not the pearls that come out on the other end, but ...'

- (205) *Monet ketjut tuntuvat olevan, -Helmiä Sioille -tyylisiä ja välillä tuntuu Sikoja syötetään Helmille...*<sup>318</sup>

lit. 'Most threads seem to be Pearls to Pigs style and sometimes it seems that Pigs **are fed** to Helmi...'

- (206) A: *Onkin jo ollut ikävä sitä sianmaksakuutioista ja perunoista tehtyä keittoa.*

B: *No eikös teidän perheessä ole ollut tapana **tarjota** sikoja Helmille?*<sup>319</sup>

lit. 'A: (I) have already been missing this soup made of pig liver cubes and potatoes.

B: Well, hasn't it been customary in your family **to offer** pigs to Helmi?'

In (207) – (209) below the same idea of pearls being fed to pigs is expressed by the presence of such elements as *huonosti sulava ruoka* 'poorly digestible food', *sianruoka* 'hogwash' and *rehu* 'forage'. Examples (208) and (209) are not tokens of the HHS constructions in a strict sense. They rather belong to another construction family – *X ETSII/ SEULOO/ POIMII/ TONKII/ LÖYTÄÄ/... HELMEN/ HELMET/ HELMIÄ ROSKASTA/ PASKASTA/ ROMUSTA* | *NP<sub>SUBJ</sub>[N{X}] Vtr{etsiä 'search'/seulaa 'sieve'/poimia 'pick'/tonkia 'dig'/löytää 'find'/...} NP<sub>OBJ</sub>{helmi 'pearl'} PP<sub>SEPAR</sub>[NP{roska 'garbage'/paska 'shit'/romu 'junk'}]* | lit. 'X searches/sieves/picks/digs/finds/... pearls from garbage/shit/junk', id. 'X searches/... some fine, valuable and rare entities from the bulk of useless stuff'. The appearance in their structure of lexical items semantically related to *sika* 'pig' is a result of host-donor interaction (described in Section 5.3.4.1 below) between these two construction families.

- (207) *Turha on heittää helmiä sioille. Ne ei helmien päälle ymmärrä - luulevat niitä vain huonosti sulavaksi ruuaksi.*<sup>320</sup>

lit. 'It's useless to throw pearls to pigs. They do not understand about pearls – (they) just think them to be **poorly digestible food**.'

- (208) *Saattaa myös löytää helmiä sianruoan seasta: voin sanoa, että se kohta Stanley Kramerin elokuvasta "Mieleton, mieleton maailma", missä tuuli puhalttaa hatun Culpepperin pasta ajotielle, ja paikalle osuva satunnainen autoilija ihan kurvaa kohti ja ajaa hatun lyttyyn, kuuluu mielestäni amerikkalaisen elokuvan kestävään kuvastoon.*<sup>321</sup>

<sup>318</sup> GG: sfnet.atk.sodat/msg/e3433770baa947cb, Nov 4 2004

<sup>319</sup> GG: sfnet.huuhaa/msg/c34ff4f5347a68fe, Apr 3 1997

<sup>320</sup> GG: sfnet.keskustelu.yhteiskunta/msg/b4960fff81b41db7, Nov 1 2006

<sup>321</sup> GG: sfnet.harrastus.elokuvat/msg/43ae7ffc753d80cc, Nov 28 2005



lit. 'You can also find pearls amongst **hogwash**: I can say that the scene from Stanley Kramer's film "It's a Mad, Mad, Mad, Mad World" where the wind blows off the hat from Culpepper's head onto the road and a motorist who happened to be on the spot just curves towards (it), drives on the hat and crushes it in my view belongs to the long-lasting catalogue of American motion picture.'

- (209) *Joten teatterikatselu jää ainoastaan harvojen helmien poimiseksi sioille tarkoitettun rehun joukosta.*<sup>322</sup>

lit. 'So theatre-going has become picking occasional pearls from **the forage** meant for pigs.'

Examples (210) and (211) below feature constructions which belong to the same family as those presented in (208) and (209) above. However, their lexical items *sikolätti* 'pigsty' and *sikalauma* 'herd of pigs' have a different semantic relation to *sika* 'pig': for the former it is of the same nature as for the item *sikala* 'piggery', which appears in (197) above; for the latter it is other kind  $F(\text{group\_of\_pigs}(a)) - \bullet(a)$ .

- (210) *Tällaisten [sosialistien] (mm. aitojen vasemmistoliberaalien) löytäminen onkin todellista helmien etsimistä sikolätistä.*<sup>323</sup>

lit. 'Finding [socialists] like this (genuine left-wing liberals among others) is truly looking for pearls in **a pigsty**.'

- (211) *Mutta pointtina on se, että niitä helmiä on baareissa niin vähän, ettei jaksa sitä sikalaumaa ensin läpi kahlata.*<sup>324</sup>

lit. 'But the point is that there are so few pearls in bars, that one can't be bothered to plough through this **herd of pigs** first.'

The majority of non-default items detected in the Google Groups-data feature some kind of semantic relation to the lexical item *sika* 'pig'. In the following two examples, NP constituents in PP<sub>ELA</sub> (212) and PP<sub>ALL</sub> (213) correspond to the words *helminauha* 'string of pearls' and *kaunis nainen* 'beautiful woman', which are semantically related to the default lexical item *helmi* 'pearl': the former via the Formal quale and the latter via the Telic quale:

- (212) *Aux- cd- tuner- tape- md- dat- vcr- liitännät ovat ns. "linjatasoisia sisäänmenoja" jotka eivät ymmärrä hienosta vinyylisoundista enempää kun sika helminauhasta.*<sup>325</sup>

<sup>322</sup> GG: sfnet.harrastus.elokuvat/msg/4c18525f52f10d55, May 6 1992

<sup>323</sup> GG: sfnet.keskustelu.libertarismi/msg/d7ef61eea434e1fa, Nov 4 1998

<sup>324</sup> GG: sfnet.keskustelu.seksi/msg/7b3165e2f1ebaa96, May 2 1998

<sup>325</sup> GG: sfnet.harrastus.audio+video/msg/06c09f5cc5b94479, Oct 19 2002

lit. ‘Aux- cd- tuner- tape- md- dat- vcr- interfaces are so-called “line-level inputs” that do not understand about the beautiful vinyl sound any more than a pig **about a string of pearls.**’

- (213) *Kielilläpuhuminen kuuluu seurakunnassa aivan alkeisiin. Lisäsin sivuilleni näytteen ja lisätietoja. Helmiä **kauniille naiselle** ja siat pysykööt **rapakoillaan**. Kiitos!*<sup>326</sup>

lit. ‘Speaking in tongues belongs to the very basics in the congregation. I added a sample and more information to my page. Pearls **to a beautiful woman** and pigs can stay **in their puddles**. Thank you!’

### 5.3.3.3.2 Non-defaults licensed by a default verb’s LCS

In this section I will examine non-default correspondence between the syntactic constituent of the category verb and various non-default phonological structures. This correspondence is assumed to be licensed by the default constituent’s lexical conceptual structure. Figure 106 below presents the formal description of the Finnish verb *heittää* ‘throw’ using Nikanne’s (1990b, 2006) thematic feature hierarchy. The theoretical basis of this formalism is presented earlier, in Chapter 3, Section 3.2.

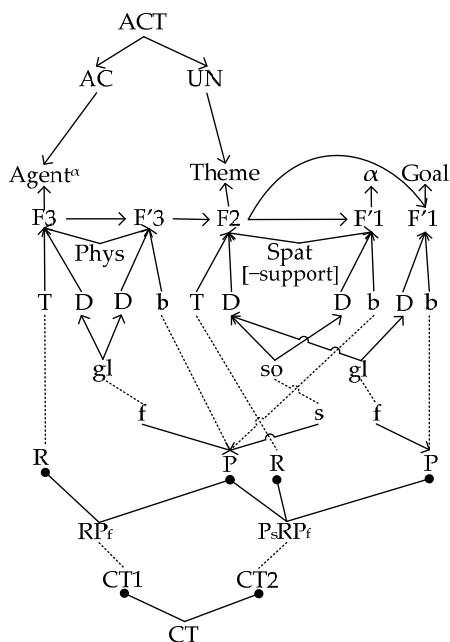


Figure 106 Formalization of the Finnish verb *HEITTÄÄ* ‘throw, cast’ with functions represented as Nikanne’s (1990b, 2006) thematic feature hierarchy

<sup>326</sup> GG: sfnet.keskustelu.uskonto.kristinusko/msg/506df6b93d081d40, Oct 12 1999

### 5.3.3.3.2.1 Causative verbs

This group includes non-default linking of the syntactic V constituent to phonological and conceptual structures of the verbs with thematic structures similar in many respects to the thematic structure of the default verb *heittää* ‘throw’. Depending on the semantic field in zones 2 and 1, causative non-defaults that occur in my data can be divided into two subgroups: verbs of caused spatial motion (5.3.3.3.2.1.1 below) and verbs of caused transfer of possession (5.3.3.3.2.1.2 below).

#### 5.3.3.3.2.1.1 Caused spatial motion

Non-default verbs of this category feature a Physical semantic field in zone 3 and a Spatial field in zones 2 and 1. Verbs in (214) – (218) below fall into the same larger category of verbs of propulsion (Jackendoff 1993: 45). They all conceptualize a similar situation, which can be characterized as ‘caused motion through the air or free space’. Thus, the spatial semantic feature [–support], which appears in the S-tier of the default verb *heittää* ‘throw’, is preserved here. The group includes such verbs as *nakata* ‘toss, chuck, bung’ (214), *viskata* ‘sling’ (215), *heitellä* ‘throw about, toss’ (216), *viskellä* ‘throw about, toss’ (217) and *ryöpyttää* ‘whirl, make gush’ (218).

- (214) [...] *ennakkovaikutelmat tulevat koetun pohjalta, enkä ole vakuuttunut vielä japanilaisen sarjakuvan/animaation ihanuudesta, mutta jos voit heittää mielestäsi joitakin helmiä, niin **nakkaappa** tälle karjulle jokunen...*<sup>327</sup>

lit. ‘[...] preconceptions come on the basis of experience, and I am not yet convinced of the loveliness of Japanese comics/animation, but if you think you can throw some pearls, then **toss** a few to this boar...’

- (215) *Tai olisi kyllä PALJON sanottavaa, mutta mietin kannattaisiko helmiä sioille **viskata**.*<sup>328</sup>

lit. ‘Or there is really MUCH to say, but I wonder if it would be worthwhile to **toss** pearls to pigs.’

- (216) *Niin, minä ja lev ja Ojala ja monet muut asialliset kreationistit olemme todenneet eräiden evolutionistien olevan täysin kyvyttömiä asialliseen keskusteluun, emmekä viitsi enää **heitellä** helmiä sioille...*<sup>329</sup>

<sup>327</sup> GG: sfnet.harrastus.kulttuuri.sarjakuvat/msg/fe50541c78c3f5ed, May 11 1998

<sup>328</sup> GG: sfnet.huuhaa/msg/a712281351ab0b41, Nov 30 2004

<sup>329</sup> GG: sfnet.keskustelu.uskonto.kristinusko/msg/4991066df2a5e028, Mar 24 2003

lit. 'So, me and lev and Ojala and many other objective creationists have found that some evolutionists are completely incapable of objective conversation, and we no longer bother to **toss** pearls to pigs...'

- (217) *Poika ottaa nimekseen Ainut Vain [...] ja lähtee aasin tammalla Jerusalemia kohti viskellen siinä ohessa helmiä sioille ...*<sup>330</sup>

lit. 'The boy takes a name of The Only One [...] and leaves on a donkey mare towards Jerusalem **tossing** pearls to pigs on his way ...'

- (218) *Kirkossa ja seurakunnissa on kahdenlaista porukkaa. Niitä kenellä on helmiä, jotka tietävät olla antamatta sioille mitään. Ja niitä jotka **ryöpyttävät** sioille kaiken aikaa jotain, mutta takuulla eivät helmiä. Sitten on ne onnettomat väliinpuotoajat jotka tulevat sikojen tallomiksi esitettyään helmensä väärässä paikassa.*<sup>331</sup>

lit. 'There are two kinds of folk in the church and in congregations. Those who have pearls, who understand not to give anything to pigs. And those who **whirl** something to pigs all the time, but definitely not pearls. Then there are those unlucky losers who get repeatedly trampled on by pigs after having shown their pearls in the wrong place.'

The peculiarity of the above verbs in (214) – (217) above can be formulated as the Agent's non-serious attitude towards the caused Event. Its formalization would require the introduction of a whole new module of representation, which would encode lexicalized mental states and attitudes (and probably much more). This is an entire research area of its own and for obvious reasons cannot be investigated within the scope of this book. A temporary ad hoc solution is presented in Figure 107 below. Jackendoff (1992: 45) follows Marr & Vaina (1982) and presents a different kind of solution for distinguishing between *throwing* and *tossing*, which involves including the 3D model representation of the action in question in the lexical entries of these verbs.

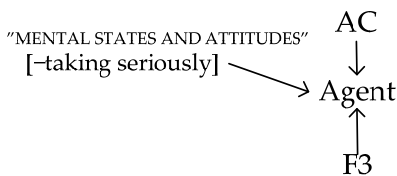


Figure 107 Attitude of the Agent in the verbs NAKATA 'toss, chuck, bung', VISKATA 'sling', VISKOA 'throw about, toss', HEITELLÄ 'throw about, toss', VISKELLÄ 'throw about, toss'

<sup>330</sup> GG: sfnet.keskustelu.uskonto/msg/e0a8775f77c6644d, Jan 17 1998

<sup>331</sup> GG: sfnet.keskustelu.uskonto.kristinuskko/msg/3b2f3302ce035c93, Oct 1 2001

In addition, the verbs *heitellä* ‘throw about, toss’ and *viskellä* ‘throw about, toss’ include a sense of iteration. Finnish frequentative verbs have been mentioned in e.g. Itkonen (1966), where they are considered to express repeated, frequently happening events. They were also studied by Wiik (1975), who suggests that the essence of the meaning of frequentative verbs is the “random course of happening”, which means that the event is non-homogeneous in time and that the various degrees of intensity as well as the periods when the event is taking place and not taking place alternate in time in a random fashion. More importantly, Wiik (1975) points out that the meaning of frequentative verbs often includes the speaker’s dismissive, disdainful attitude and playfulness. Derivation mechanisms of Finnish frequentatives are also briefly described in Ojanen & Uotila-Arcelli (1979) and in more depth in Suihkonen (1994), who examines frequentative derivation in Finnish dialects. According to Suihkonen (1994), the derived verbs containing the *-ele-* infix are the dominant type in all Finnish dialects. Although Suihkonen (1994) does not study frequentatives from the semantic point of view and concentrates entirely on morphology, his book includes a brief introduction into the concept of frequentativity.

Jackendoff (1990) reflects on a possible way to encode the sense of iteration in a conceptual structure. His proposal is to regard it as an operator, which maps a conceptual constituent that encodes a single Event into a conceptual constituent that encodes a repeated sequence of individual Events of the same type. He also remarks that this operator has exactly the same semantic value as the plural marker, which maps a conceptual constituent that encodes an individual Thing into a conceptual constituent that encodes a collection of Things of the same type (Jackendoff 1990: 29). Following Jackendoff, I introduce a plural marker into the formal representation of the Finnish verbs *heitellä* and *viskellä* ‘throw about, toss’ in Figure 108 below. It is presented in the form *> ONE*, i.e. ‘more than one’. The tier, in which the marker can be located, is probably the one that includes all kinds of quantifying markers (some suggestions on the topic can be found e.g. in Larjavaara (2007: 193-194), but it is not the purpose of this book to embark into this area. A possible structure of the quantifying tier can be constructed following the lanes of the traditional distinction between [*count*] and [*uncount*] features. The former category can include numerals (ONE, TWO, THREE, ...) and such

operators as EVERY, EACH, SEVERAL (> ONE), etc., and the latter such operators as MUCH and LITTLE.

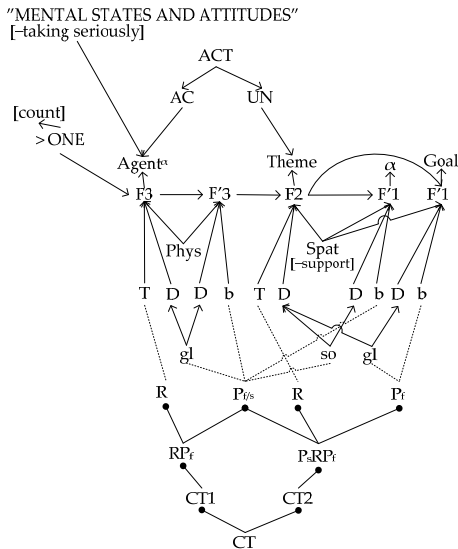


Figure 108 HEITELLÄ, VISKELLÄ 'throw about, toss'

The formal representation of the verb *ryöpyttää* 'whirl, make gush' in Figure 109 includes an additional quantifying marker MUCH, which selects the Theme. This is to indicate the large quantity of this argument (cf. the meaning of the word *ryöppy* 'spate').

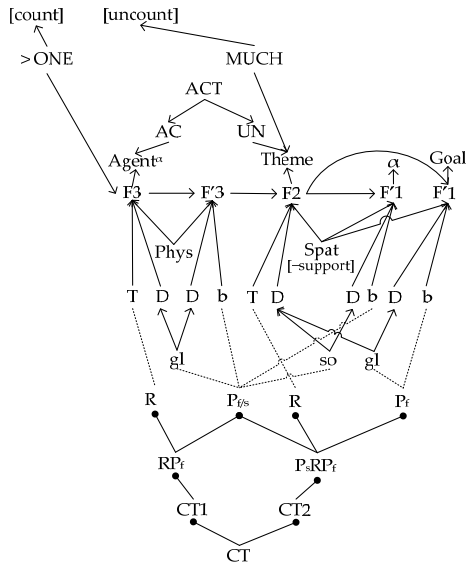


Figure 109 RYÖPYTTÄÄ 'whirl, make gush'

Both the compound verb *pistää tulemaan* ‘let come, set coming’ in (219) below and *tuoda* ‘bring’ in (220) still belong to the ‘caused spatial motion’ category, but as opposed to the verbs in (214) – (218) above, they are more abstract in the sense that they do not specify that the caused spatial motion happens through the air or free space, hence the missing [-support] feature.

- (219) *Joten **pistä** niitä helmiäsi vaan **tulemaan** :-)*<sup>332</sup>  
lit. ‘So let those pearls of yours come :-))’

- (220) *”Kelvottomaksi” kokeen voi tehdä valmiiksi päätetty tulkinta kaikista mahdollisista esiintulevista tapahtumista, eikä sellaiseen ”loukkaukseen” varmaan Jumala halua **tuoda** ”helmiä sioille”.*<sup>333</sup>  
lit. The test can be made “useless” by the pre-decided interpretation of all possible upcoming events, and it is certainly not for an “infringement” like this that God wants to **bring** “pearls to pigs”.

Both verbs conceptualize some deictic information. Larjavaara (1990: 256f) analyzes deictic properties of the Finnish verbs *tulla* ‘come’ and *mennä* ‘go’. He admits that both verbs have an enormous number of different uses and therefore are subjects to semantic variation. Nevertheless, he presents the following base meanings of these verbs:

*tulla* = [+MOVE, +APPROACHING]  
*mennä* = [+MOVE, –APPROACHING]

A similar distinction could be made for *tuoda* ‘bring’ and *viedä* ‘take away’:

*tuoda* = [+CAUSE MOVE, +APPROACHING]  
*viedä* = [+CAUSE MOVE, –APPROACHING]

In the formal description of the compound verb *pistää tulemaan* ‘set coming’ presented in Figure 110 below and the one of the verb *tuoda* ‘bring’ in Figure 111 below there is a deictic feature [+APPROACHING], which selects the Goal argument of the thematic tier. Note that the temporal structure of *tuoda* ‘bring’ differs from that of the verbs I have been looking at so far. It represents the relation of entrainment (see Section 3.2.5 in Chapter 3 for a distinction between launching and entrainment).

<sup>332</sup> GG: sfnet.keskustelu.rakentaminen/msg/cac0ce7de004e0ac, Aug 28 2005

<sup>333</sup> GG: sfnet.keskustelu.uskonto.kristinusko/msg/5377ceea628fbee1, Apr 14 2003

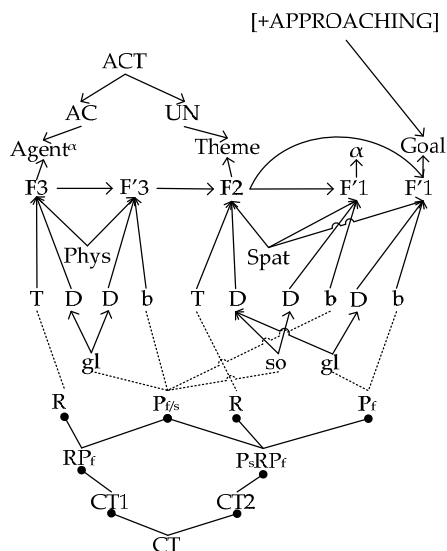


Figure 110 PISTÄÄ TULEMAAN 'set coming'

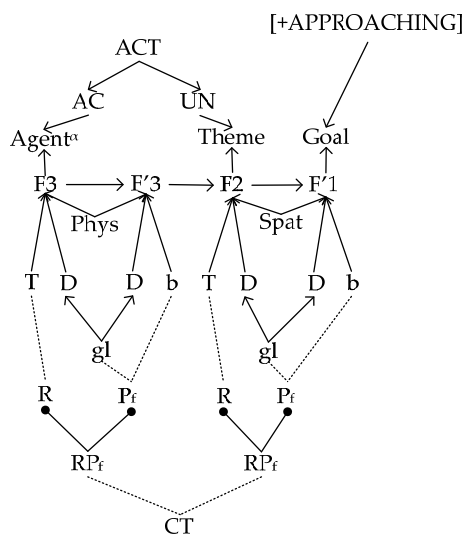


Figure 111 TUODA 'bring'

#### 5.3.3.3.2.1.2 Caused transfer of possession

In situations conceptualized by the group of verbs in the previous section the Theme undergoes a change of location. Non-default lexical items analyzed in the current section conceptualize the caused transfer of possession. Obviously, we are dealing here with conceptual parallelism, formulated in Jackendoff (1992: 62) as follows:



Change in physical location of X parallels Change in possessor of X

Lakoff (1987) and Lakoff & Turner (1989) would claim that the process underlying this parallelism is metaphorical extension, which relates spatial concepts with possessional ones. Jackendoff (1992: 60) also agrees that there is a connection between sensorimotor concepts, such as spatial and abstract ones, like possession. However, he criticizes Piaget (1966) and Lakoff (1987) for missing a crucial logical part of the process of extension: it is not possible to construct the notion of ownership from any combination of spatial primitives (Jackendoff 1993: 63). Jackendoff (1976, 1983, 1992) suggests an alternative – the Thematic Relations Hypothesis, which claims that the crucial factor in the parallelism between the conception of physical space and the notion of possession is an innate abstract organizing system for concepts, which is able to have different realizations (Jackendoff 1992: 64-65). This abstract system – the semantic field tier (S-tier) – has already been mentioned in Chapter 3, Section 3.2.6. Thus, the main distinction of non-default verbs presented below from the verbs in 5.3.3.3.2.1.1 above is the Possessive semantic field in zone 2 and zone 1. In Nikanne's (1990a) classification of possessive expressions it corresponds to Poss1, which describes alienable possession. The Causer can be assigned the role of a Donor and the Landmark the role of a Recipient.

The most general verbs of this group are *antaa* 'give' (221) and *välittää* 'provide, supply' (222):

(221) *A: Ja monet naiset kokevat suhteen teekkariin hankalaksi sen takia, että eivät pidä näitä puoleensavetävinä.*<sup>334</sup>

*B: Totta. Miksi antaa helmiä sioille.*<sup>335</sup>

lit. 'A: And many women experience a relationship with technology students as difficult because they do not consider them to be the attractive.

B: True. Why **give** pearls to pigs.'

(222) *Miten muuten ajattelit välittää nämä helmet internet-yhteyden ja tietokoneen välityksellä?*<sup>336</sup>

lit. 'By the way, how did you intend to **provide** these pearls through the Internet and computer?'

<sup>334</sup> GG: sfnet.keskustelu.seksi/msg/260c02b5f36d0854, Feb 6 1997

<sup>335</sup> GG: sfnet.keskustelu.seksi/msg/517691b01058d2da, Feb 6 1997

<sup>336</sup> GG: sfnet.keskustelu.evolutio/msg/9684fef88f3c950e, Mar 5 2003



successfulness of the causation is expressed by means of the three possible values that the causative function CAUSE may have: a successful causation is marked as CAUSE<sup>+</sup> and an unsuccessful one as CAUSE<sup>-</sup>, while CAUSE<sup>U</sup> indicates that it is not possible to determine whether the causation is successful or not. Nikanne (1990b, 2006) formalized the notion of success as the boundedness of F'3:

$$\text{CAUSE}^- = [\text{F3} [\text{T}][\text{D}] \rightarrow \text{F}'3 [\text{D} [\text{gl}]]]$$

$$\text{CAUSE}^+ = [\text{F3} [\text{T}][\text{D}] \rightarrow \text{F}'3 [\text{b}][\text{D}[\text{gl}]]]$$

$$\text{CAUSE}^U = [\text{F3} [\text{T}][\text{D}] \rightarrow \text{F}'3 <[\text{b}]>[\text{D}[\text{gl}]]]$$

Thus, the causation is successful iff it is bounded. The absence of the boundedness-feature from function F'3 means that the causation is unsuccessful. Putting the [b] feature inside angle brackets, as in Figure 113 below, indicates its optionality, which leaves the matter of successfulness unresolved.

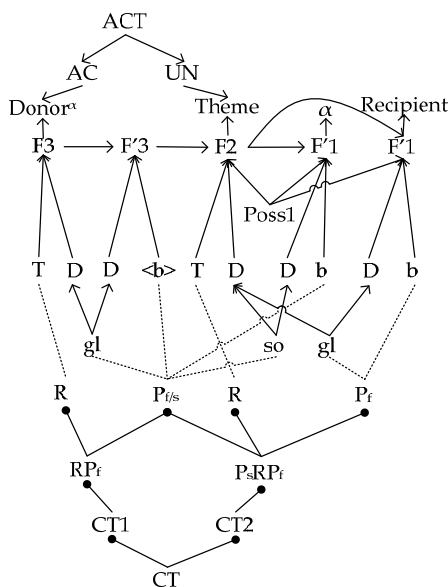


Figure 113 TARJOTA, TARJOILLA 'offer'

In (225) and (226) below two Finnish verbs *jakaa* and *jaella* both meaning 'distribute' encode that portions or shares of the Theme are repeatedly allotted to several recipients.

(225) *Mutta voi olla että jossain asiassa olen itse ainoastaan ja vain oikeassa. Se on salaisuus. En toki helmiä **jakele**. Ne on arvokkaita!*<sup>339</sup>

<sup>339</sup> GG: sfnet.keskustelu.uskonto/msg/4e31d0229d9b2a8d, Dec 7 1995

lit. 'But it may be that in something only I myself am exclusively right. It is a secret. I certainly do not **distribute** pearls. They are valuable!'

- (226) *Liitän äänitiedostoja ja kuva tiedostoja jos helmiä sioille viittii **jakaa** ja miksei.*<sup>340</sup>  
 lit. 'I will add sound files and image files if (I will) bother to **distribute** pearls to pigs and why not.'

Formal representation of these verbs in Figure 114 above include several quantifying markers: the f-chain and the recipient are selected by the plural marker and the theme is selected by the marker SOME. In addition, F'1 is selected by the thematic feature [di], i.e. 'distributed', which indicates that the Theme is distributed over the Path or Place.

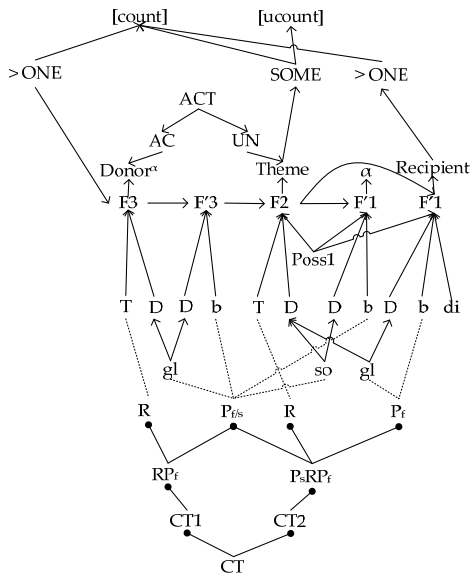


Figure 114 JAKAA, JAELLA 'distribute'

In the LCS of the previously examined possessive verbs it was not specified whether there is a legal change of ownership involved in the Event. The following verbs *lahjoittaa* 'donate' (227), *myydä* 'sell' (228)<sup>341</sup>, and *varastaa* 'steal' (229) include [+ownership] and [±legal] features in the Possessive semantic field of the S-tier (Figures 115, 116 and 117 below). *Lahjoittaa* 'donate' does not include any medium of exchange – the transition is done without any expectation or receipt of an equivalent on the part of the Donor. *Myydä* 'sell', however, includes an embedded

<sup>340</sup> GG: sfnet.atk.amiga/msg/629e63d1288f8ced, May 9 2006

<sup>341</sup> The same example is also analyzed as a referentially licensed non-default in Section 5.4.2.2 of the current chapter.

conceptual clause, which involved the transition of the ‘Medium of Exchange’ (the intuitive interpretation of this concept would be ‘Money’, but in our era of electronic transactions it is not so clear anymore). Cf. Jackendoff’s (1983, 1990) analysis of transactional verbs, such as *buy* and *sell*, which includes two conceptual clauses – the primary clause that conceptualizes the transfer of goods and the subordinate clause that represents the transfer of money. Jackendoff’s (1983, 1990) description of the subordinate clause includes the thematic function EXCH, while my solution presented in Figure 116 below lacks it. Instead, the MEDIUM OF EXCHANGE argument indicates that the goods are exchanged for money.

(227) A: *Puu on paljon arvokkaampaa pidemmälle jalostettuna kuin poltettuna.*

B: *Niinpä, vaan helmeä EI SAA LAHJOITTAA sikaileville valtion eliittimonopoleille.*<sup>342</sup>

lit. ‘A: Wood is far more valuable when it is further refined than when it is burnt

B: Sure, one SHOULD NOT **DONATE** a pearl to swinish elite state monopolies.’

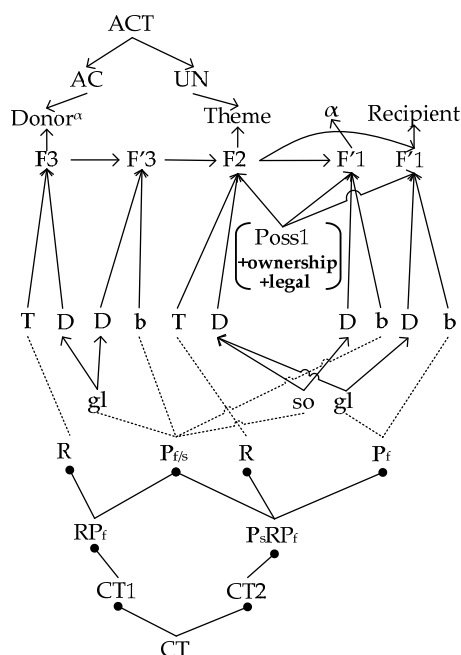


Figure 115 LAHJOITTAA ‘donate’

<sup>342</sup> GG: sfnet.keskustelu.energia/msg/0d662885fd59d4bd, Aug 1 2005

(228) Miksi *myydä* helmiä sioille...vai pitäisikö sanoa possuille.<sup>343</sup>

'Why **sell** pearls to pigs ... or should I say to piggies.'

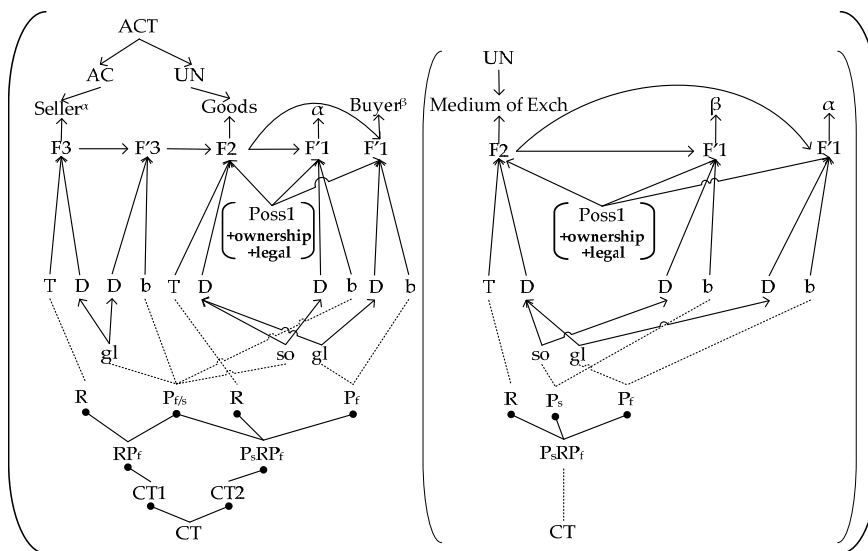


Figure 116 MYYDÄ 'sell'

(229) Varastakaa helmiä sioilta!<sup>344</sup>

'Steal pearls from pigs!'

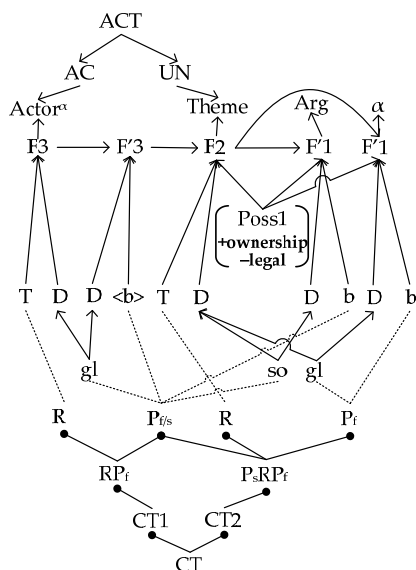


Figure 117 VARASTAA 'steal'

<sup>343</sup> GG: sfnet.harrastus.autot/msg/599b744db8a1390c, Feb 28 2002

<sup>344</sup> GG: sfnet.urheilu.jaakiekko/msg/c988936be180a5c4, Nov 11 1996

The verb *tyrkyttää* ‘impose’ in (230) below describes a situation where the Recipient is not willing to accept the Theme. It can be formalized by assigning the ACT role Undergoer to this thematic argument and by adding a feature [–volitional], or [–vol] (Figure 118 below). The action tier is described in more detail in Section 3.2.4 of Chapter 3.

(230) *Tietenkin ongelma saattaa silloin olla se, että olet tyrkyttämässä helmiä sioille.*<sup>345</sup>

lit. ‘Of course, then the problem may be that you are trying to **impose** pearls on pigs.’

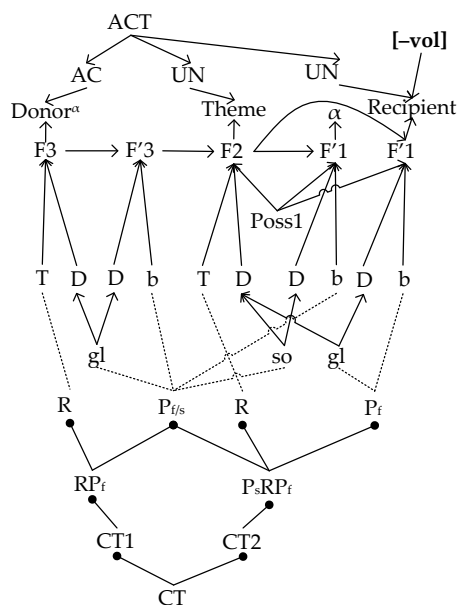


Figure 118 TYRKYTTÄÄ ‘impose’

### 5.3.3.3.2.2 Intransitive spatial motion verbs

According to Jackendoff (1992: 55) there are relations among conceptual structures that specify how to pass from one concept to another. These relations are called inference rules. Situations encoded by intransitive verbs presented in this section are licensed by the following inference rule (Jackendoff 1990: 39):

$x \text{ cause } E \text{ to occur} \rightarrow E \text{ occur}$

<sup>345</sup> GG: sfnet.keskustelu.psychologia/msg/5ede67f00993a5ca, Sep 14 2001

In Nikanne (1990b: 120) it is reformulated as presented in Figure 119 below:

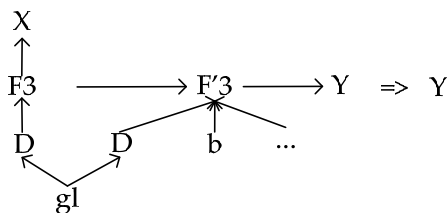


Figure 119 Inference rule '[CAUSE<sup>+</sup> [(X), (Y)]] => Y' reformulated in Nikanne (1990b: 120)

The inference rule which licenses non-default non-causative constructions in the *HHS* construction family is shown in Figure 120 below.

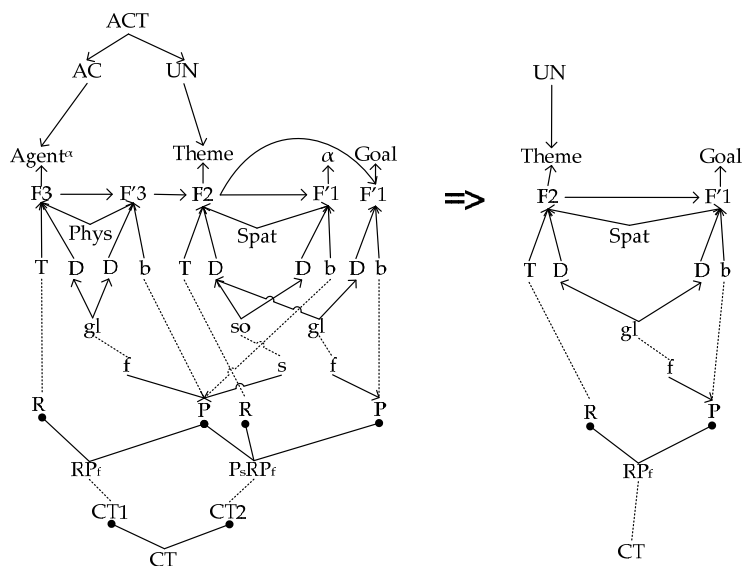


Figure 120 Inference for HEITTÄÄ 'throw'

The verb *lentää* 'fly' in (231) below (Figure 121) can be regarded as the most closely related inference of the default *heittää* 'throw'. It encodes motion through air or free space and thus contains, just like the default does, the semantic feature [-support] in the Spatial field.

- (231) *Taas lensivät Tean satiirin helmet sioille...*<sup>346</sup>  
lit. 'Again pearls of Tea's satire **flew** to pigs.'

<sup>346</sup> GG: sfnet.huuhaa/msg/53773e422c631c8a, Oct 9 1997



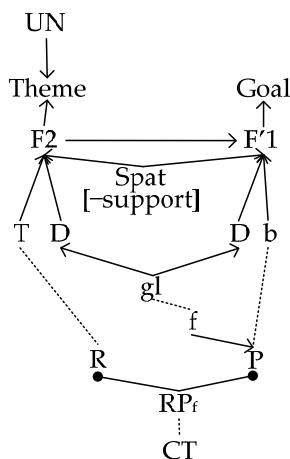


Figure 121 LENTÄÄ ‘fly’

Another ‘motion through air’ verb is *ryöpsähtää* ‘shower, gush (out), blast, flood’ (232). It is a momentative verb encoding a single burst of some large quantity (cf. the causative *ryöpyttää* ‘whirl’ in (218) above formalized in Figure 109).

- (232) *Nyt jätkät ja gimmat äkkiä keräämään: taas **ryöpsähti** läjä helmiä sioille kun nimimerkki "J.Alkula" u...@hotmail.com näytti persoonakohtaisesti esimerkkiä, miten urheasti ja esimerkillisesti kielipoliittistakin keskustelua voidaan käydä nimimerkin suojasta, vihjailemalla, hipaisemalla ja kertomalla, että perusasioiden hallinta omassa elämässä ei ole ollenkaan haitaksi.*<sup>347</sup>

lit. ‘Come on guys and gals and quickly gather: a pile of pearls to pigs **gushed out** again when the username "J.Alkula" u...@hotmail.com showed a personal example of how bravely and exemplarily one can carry on even a language-political debate under the protection of a nickname, by implying, touching upon and telling that mastering the basics in one’s own life is not at all detrimental.’

The following pair of deictic intransitive spatial motion verbs includes *mennä* ‘go’ (233) and *tulla* ‘come’ (233). Conceptual structures of these verbs, formalized in Figure 122 and Figure 123 below, are otherwise identical, except for the deictic feature [ $\pm$ APPROACHING], which selects the Goal argument from the thematic tier.

- (233) *Vosin ladata sulle kilometreittäin faktaa asiasta, mutta se lienee täysin hyödytöntä [...] **Menisi** niin sanotusti helmiä sioille.*<sup>348</sup>

<sup>347</sup> GG: sfnet.keskustelu.kielipoliitikka/msg/34a51ebbec7c16c8, Sep 25 2001

<sup>348</sup> GG: sfnet.keskustelu.uskonto/msg/2d83306495eabf31, Jul 11 1996

lit. 'I could load you with kilometers of facts on this matter, but it is probably completely pointless [...] Pearls would, so to speak, **go** to pigs.'

- (234) *Nyt tuli helmiä sialle, minä ymmärrän runoutta paremmin sitä matematiikan kieltä.*<sup>349</sup>

lit. 'Now pearls **came** to a pig, I understand poetry better than this language of mathematics.'

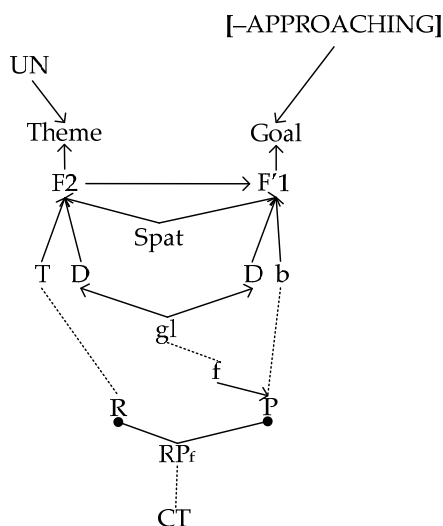


Figure 122 MENNÄ 'go'

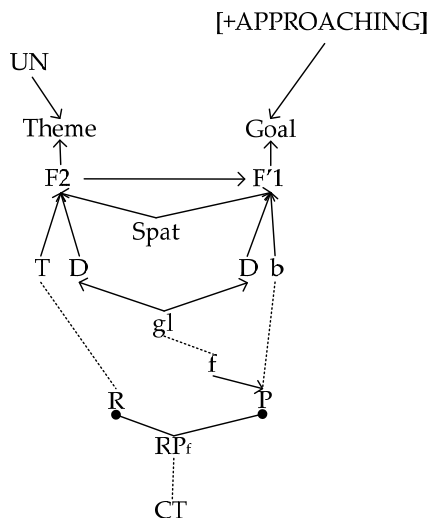


Figure 123 TULLA 'come'

<sup>349</sup> GG: sfnet.keskustelu.ihmissuhteet/msg/8d4ff6f29ec31de7, Oct 28 2003

### 5.3.3.2.3 Stative location and possession

Just like the intransitive spatial motion situations in the previous section were inferred from the caused motion situation, conceptual structures presented here are inferred from the intransitive situation. An inference rule formalized in Figure 124 below was formulated in Jackendoff (1987) and mentioned in Nikanne (1990b: 181).

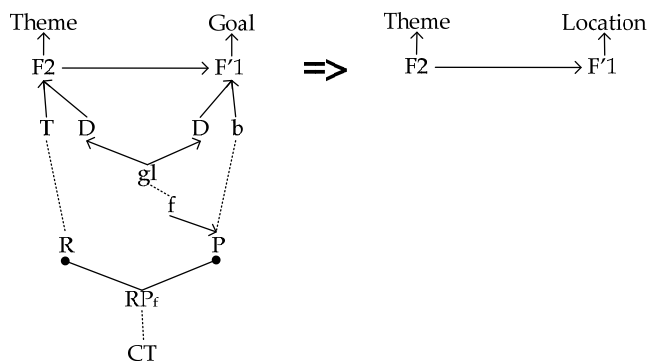


Figure 124 Inference rule for  $BE \rightarrow AT$

Here one can find both a spatial situation conceptualized by the verb *löytyä* 'be found, occur' in (235) and a possessive construction  $NP_{ADE}\{X\}$  on *Y* 'at *X* is *Y*; *X* has *Y*' in (236) represented in Figure 125 and Figure 126 below.

- (235) *Totta on että monissa marketeissa laarit ovat täynnä pa\*kaa mutta ainakin Jyväskylän Gigantista löytyy useimmiten helmiä sioille...*<sup>350</sup>  
lit. 'It is true that in many supermarkets bins are full of sh\*t but at least in Jyväskylä's Gigantti pearls to pigs can **be found** in most cases ...'
- (236) *Soini, voi olla helmi siallakin.*<sup>351</sup>  
lit. 'Soini, even a pig can **have** a pearl.'

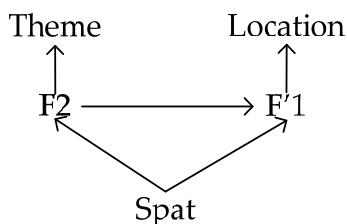


Figure 125 LÖYTYÄ 'be found, occur'

<sup>350</sup> GG: sfnet.harrastus.pelit/msg/fde3c9af3e8744bb, Feb 19 2004

<sup>351</sup> GG: sfnet.harrastus.mp/msg/5bbc5558918a788f, Sep 9 2003

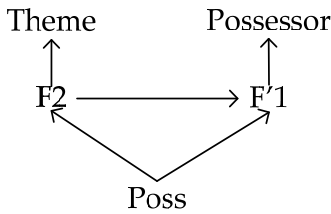


Figure 126 OLLA 'be'

#### 5.3.3.4 Non-default lexical linking licensed by the CS/PU

The CS/PU can also license non-default lexical items. In my data, at least two types of such licensing have been detected. The first group includes both transitive (*tuhlata* 'waste, squander, fritter away' (239), *uhrata* 'sacrifice' (240), *panna likoon* 'stake' (241), *syytää* 'ladle out' (242)) and intransitive (*joutua* 'end up, finish up, fall (in)to' (237), *lipsahtaa* 'slip' (238)) verbs all having one thing in common: the situations which are conceptualized by them are 'unintended', 'unwanted', 'accidental', 'unfavourable', 'useless' and 'inexpedient'. These modal features can in fact be traced back to the *PrNess*→*Neg* modal features in the CS/PU, previously described in Chapter 3, Section 3.4.3.6.2 and Section 5.3.2 of the current chapter.

- (237) JR:n sanat sopivat oikeinkin hyvin luonnehtimaan tämän threadin aloittanutta viestiä, tuota ehkä tahatonta, ehkä tahallista (tai sekä että) superraskaan pottuilun(PC) helmeä, joka joutui sioille, kuten pitikin.<sup>352</sup>

lit. 'JR's words fit very well to describe the message that started this thread, that maybe unintentional, maybe intentional (or both) (PC) pearl of super heavy aggravation, which ended up to pigs, as it should.'

- (238) A: Miten nuo helmet liittyy tähän viestiin?  
B: Taisivat **lipsahtaa** possujen kaukaloon :-)<sup>353</sup>

lit. 'A: How are these pearls related to this message?

B: They seemed to slip into the piggies' trough :-)'

- (239) En viitsi **tuhlata** helmiä sioille.<sup>354</sup>

lit. 'I can't be bothered to **waste** pearls on pigs.'

- (240) A: Sinulla ei kuitenkaan ole mitään rationaalisia vasta-argumentteja.

B: Paljonkin, mutta keskustelutaktiikkasi tuntien, en viitsi **uhrata** "helmiä sioille".<sup>355</sup>

<sup>352</sup> GG: sfnet.keskustelu/msg/992c945d72bd26fa, May 23 1997

<sup>353</sup> GG: sfnet.tori.myydaan.atk.komponentit/msg/b3880fa624e6109d, Feb 23 2003

<sup>354</sup> GG: sfnet.keskustelu.uskonottomuus/msg/f7a1664f6b298148, Feb 4 2003

lit. 'A: You do not, however, have any rational counter-arguments.

B: Actually (I do have) a lot, but knowing your conversation tactics, I can't be bothered to **sacrifice** "pearls to pigs".'

- (241) *Minä olen tästäkin aiheesta esittänyt ihan pätevän analyysin, mutta kaiken maailman pöhköt ei sellaisia kykene ymmärtämään. Pannaan nyt vielä kerran likoon pari "helmeä sioille".*<sup>356</sup>

lit. 'I have presented quite a valid analysis on this subject as well, but all sorts of blockheads are incapable of understanding things like that. Let's just **stake** one more time a couple of "pearls to pigs".'

- (242) *Älä pliis tuhlaa aikaasi ja nyytisten kaistaa **syytääksesi** helmiä sialle.*<sup>357</sup>

lit. 'Please, do not waste your time and news bandwidth on **ladling out** pearls to a pig.'

The second group contains such verbs as *ansaita* 'deserve' (242) and *kelvata* 'fit, suit, be good enough' (243). The licensing source here is the property INADEQUATE presented in Figure 64 (Chapter 3, Section 3.4.3.5).

- (243) *Käsittelen nyt sitä Tiede-lehden numeroa 8/2006, vaikka tuollaiset evosiat ei tällaisia helmiä **ansaitsekaan**...*<sup>358</sup>

lit. 'I will now discuss the Science magazine issue 8/2006, although evo-pigs like that do not **deserve** such pearls ...'

- (244) *Helmet eivät **kelpaa** ravinnoksi sioille, ei vaikka kuinka kimaltelisivat. Ja päinvastoin.*<sup>359</sup>

lit. 'Pearls are **unfit** for pigs' nutrition, no matter how they shine. And vice versa.'

It should be mentioned that linking to syntax in these examples does not follow the default pattern. In (242) the  $NP_{PL}\{evosika \text{ 'evo-pig'}\}$ , which by default is a part of the  $PP_{ADJUNCT}$ , is linked to the subject position, while in (243) the subject argument corresponds to the  $NP_{PL}\{helmi \text{ 'pearl'}\}$ , which in the default case would be linked to the object.

<sup>355</sup> GG: sfnet.keskustelu.politiikka/msg/82f00e3ce5dfe8c0, May 6 2006

<sup>356</sup> GG: sfnet.keskustelu.talous/msg/e052705ad5921c08, Dec 3 2006

<sup>357</sup> GG: sfnet.harrastus.elektroniikka/msg/a8778411e852840b, Oct 9 2006

<sup>358</sup> GG: sfnet.keskustelu.evolutio/msg/d9464d2300559157, Dec 14 2006

<sup>359</sup> GG: sfnet.keskustelu.kielipolitiikka/msg/c855b46a87a90173, Mar 5 2003

### 5.3.4 Interaction between constructions of different construction families

The phenomenon discussed in this section is largely known by the term contamination defined as “a combination of parts of two or more PUs which may originate in a new PU” (Melerovič & Mokienko 2001); or phraseological blending defined as “formal structural fusions of two phraseological units into one modification, in which the resulting unit shares one or more lexemes from input units” (Omazič 2007). The former term contains negative connotations, while the latter makes a strong reference to the theory of Blending and Conceptual Integration, reviewed in Section 4.3 of Chapter 4. For these reasons I prefer to talk about interaction between constructions of different families. In Sections 5.3.4.1 and 5.3.4.2 I will discuss different types of such interaction – host-donor interaction, embedding and coordination.

#### 5.3.4.1 Host-donor interaction

To understand the underlying mechanisms of PU host-donor interaction, one has to look at the structural and semantic properties of the interacting units. Let us look at the Examples in (245) – (249) below, where the interacting parties are, on the one hand, constructions from the *HHS* construction family, and, on the other hand, constructions from the *KSH* construction family (see Section 5.3.3.1 for a more detailed discussion on this particular construction). The latter is represented by:

- a. The negative construction *Y EI YMMÄRRÄ/TIEDÄ Z:STA ENEMPÄÄ KUIN SIKÄ HOPEALUSIKASTA* | *NP<sub>SUBJ</sub>[N{Y}] NEG{ei ‘not’} Vcogn{ymmärtää ‘understand’/tietää ‘know’} PP<sub>ELA</sub>[NP{Z}] COMP{enempää kuin ‘more than’} NP<sub>SUBJ</sub>{sikä ‘pig’} PP<sub>ELA</sub>[NP{hopealusikka ‘silver spoon’}]* | lit. ‘Y does not understand Z/know about Z more than a pig about a silver spoon’, id. ‘Y does not understand Z/know about Z at all’.
- b. The affirmative construction *Y YMMÄRTÄÄ/TIETÄÄ Z:STA YHTÄ PALJON KUIN SIKÄ HOPEALUSIKASTA* | *NP<sub>SUBJ</sub>[N{Y}] Vcogn{ymmärtää ‘understand’/tietää ‘know’} PP<sub>ELA</sub>[NP{Z}] COMP {yhtä paljon kuin ‘as much as’} NP<sub>SUBJ</sub>{sikä ‘pig’} PP<sub>ELA</sub>[NP{hopealusikka ‘silver spoon’}]* | lit. ‘Y understands Z/knows about Z as much as a pig about a silver spoon’, id. ‘Y does not understand Z/know about Z at all’.

In the resulting unit, the structure is inherited either from (a) like in (245) – (247) below, or from (b) like in (248) and (249) below, but in the resulting structure the noun constituent in the *NP<sub>ELA</sub>* corresponds to *helmi* ‘pearl’ instead of *hopealusikka* ‘silver spoon’. I will call the interacting

construction which retains its syntactic structure the host construction, and the construction which provides it with a lexical element the donor construction.

- (245) *Etenkin Sakari Kinnunen ja nimimerkki VanillaQueen/Tatya ovat oikein kunnostautuneet tässä ryhmässä skeptikoiden haukkumiseksi ulkona oleviksi äärimmäisen vittumaisiksi ihmisiksi jotka eivät ymmärrä ylempien värähtelytasojen sfääreistä enempää kuin sika helmistä.*<sup>360</sup>

lit. 'In particular, Sakari Kinnunen and the username VanillaQueen/Tatya have really distinguished themselves in this group as being extremely unpleasant human beings who are out there to denigrate skeptics and who **do not understand about** the spheres of the higher vibration levels **any more than a pig about pearls.**'

- (246) *Kannattaako vääntää howtoa, jos ei ymmärrä Ohjelmasta senkään vertaa mitä siat helmistä? ;)*<sup>361</sup>

lit. 'Is it worth sifting through the how-to if you **don't understand about** the Program **even as much as pigs about pearls?**'

- (247) *Aux- cd- tuner- tape- md- dat- vcr- liitännät ovat ns. "linjatasoisia sisäänmenoja" jotka eivät ymmärrä hienosta vinyylisoundista enempää kun sika helminauhasta.*<sup>362</sup>

lit. 'Aux- cd- tuner- tape- md- dat- vcr- interfaces are the so called "line-level inputs" that **do not understand about** the beautiful vinyl sound **more than a pig about a string of pearls.**'

- (248) *Havaitsemme, että tiedät sosialismista likimain yhtä paljon kuin sika helmistä.*<sup>363</sup>

lit. 'We see that you **know about** socialism nearly **as much as a pig about pearls.**'

- (249) *Tiedätte epilepsiasta yhtä paljoa kuin sika helmistä!*<sup>364</sup>

lit. 'You **know as much about** epilepsy **as a pig about pearls!**'

Figure 127 below demonstrates the intersection of both constructions (overlapping elements are marked with solid grey colour). Again, the LCSs of the lexical items *helmi* 'pearl', *sika* 'pig' and *hopealusikka* 'silver spoon' do not contain the complete semantic descriptions of the items in question, but only the evaluative elements.

<sup>360</sup> GG: sfnet.keskustelu.rajatieteet/msg/9c2b1f8cd437ea17, May 28 2003

<sup>361</sup> GG: sfnet.atk.linux/msg/567d6608290ded1d, Mar 3 2002

<sup>362</sup> GG: sfnet.harrastus.audio+video/msg/06c09f5cc5b94479, Oct 19 2002

<sup>363</sup> GG: sfnet.keskustelu.politiikka/msg/e5666e19d295d897, Jun 17 1998

<sup>364</sup> GG: sfnet.keskustelu.varaventiili/msg/8a6c3be4b14334d5, Sep 9 2005

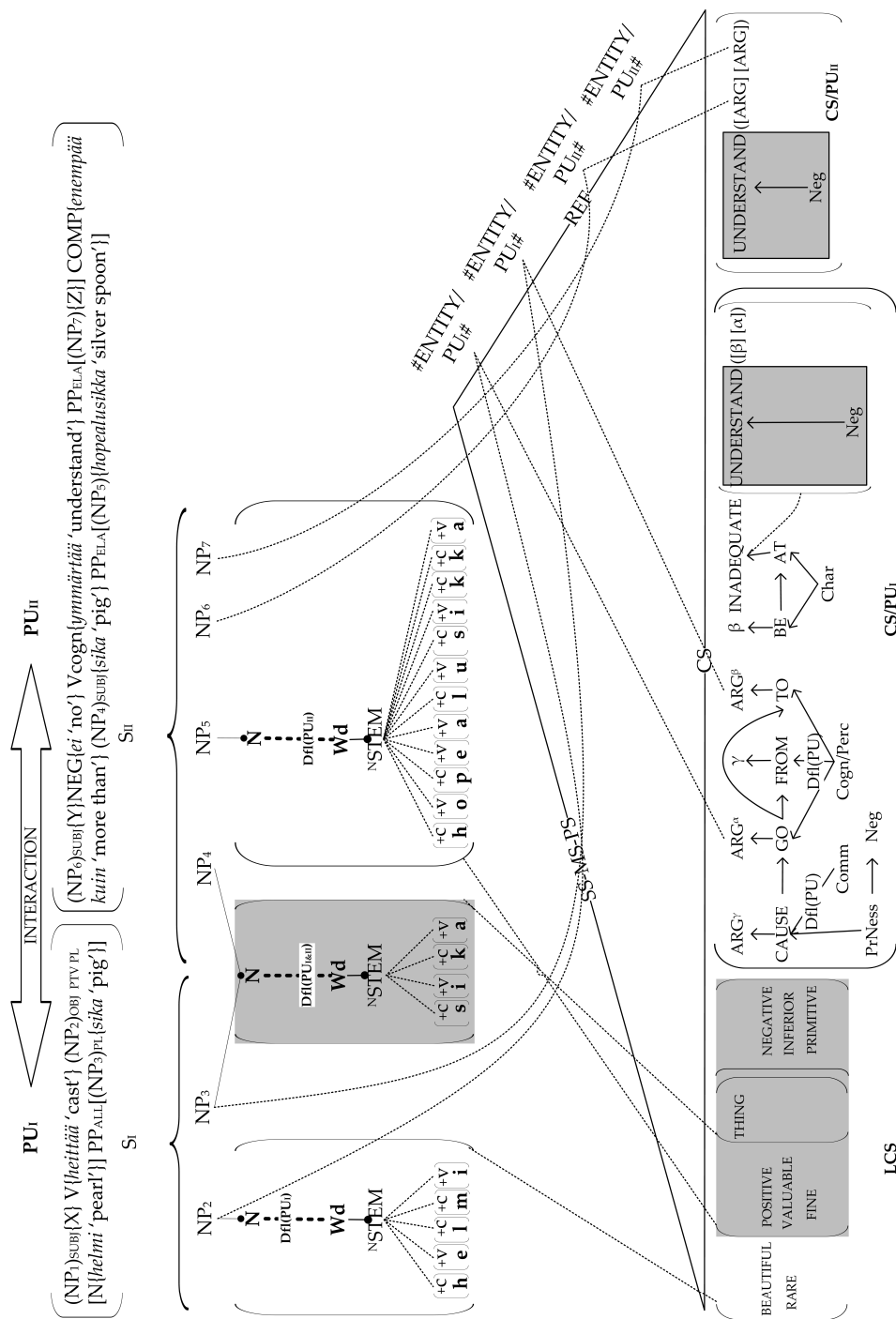


Figure 127 Structural similarities in the interacting constructions



In Figure 127, the constituent N{sika ‘pig’} is the element which both constructions share on the formal level. In both constructions the conceptual structure of this lexical item activates on the Evaluative level<sup>365</sup> such properties as NEGATIVE, INFERIOR and PRIMITIVE (institutionalized evaluative properties like DIRTY, FAT, GROSS, RAVENOUS, which are also associated with the concept of pig are not activated in these particular constructions). In addition, both involve a contraposition with another entity of the category THING (*helmi* ‘pearl’ and *hopealusikka* ‘silver spoon’), which on the level of institutionalized associations activates the feature FINE and on the level of social status and value the features POSITIVE and VALUABLE<sup>366</sup>. Thus, both constructions imply that a primitive entity (a pig in both cases) cannot comprehend a fine entity. A similar contraposition is also present in the following Russian (250), (251) and Finnish (252) PUs:

- (250) ЗНАЕТ ТОЛК, КАК СВИНЬЯ В АПЕЛЬСИНАХ (Dal’ 2000)  
 [Znáet tolk, kak svin’já v apel’sínah]  
 lit. ‘(Somebody) is an expert (in something) like a pig in oranges’  
 id. ‘Somebody does not know/understand something at all’
- (251) НЕ СВИНЫМ РЫЛОМ ЛИМОНЫ НЮХАТЬ (Dal’ 2000)  
 [Ne sviným rýlom limóny n’úhat’]  
 lit. ‘A snout isn’t good enough for smelling lemons’  
 id. ‘Somebody is not good enough to do something’
- (252) HULLU EI HYVVEE TUNNE, EIKÄ SIKÄ SUOLOO TAJUVA<sup>367</sup>  
 lit. ‘An idiot does not know good, nor does a pig appreciate salt’  
 id. ‘Somebody cannot appreciate something good’

Given that the constructions share the above-mentioned structural and semantic properties, their interaction is accomplished in the following way:

- The shared lexical item *sika* ‘pig’ is preserved in both constructions.
- The interchanged lexical items are *helmi* ‘pearl’ and *hopealusikka* ‘silver spoon’.

<sup>365</sup> Evaluation as a part of LCS of nouns has been discussed in Section 3.4.1.3.2 of Chapter 3.

<sup>366</sup> In the Finnish and English languages silver spoon is associated with wealth; someone born into a wealthy family is said to ‘be born with a silver spoon in (his) mouth’ *SYNTYÄ HOPEALUSIKKA SUUSSA*.

<sup>367</sup> Sananparsikokoelma (2010) - Riistavesi, S. Heikkinen, 1932

- Lexical items are interchanged depending on which construction functions as a host, and which as a donor. The host construction borrows the interchangeable item from the donor construction while preserving its morphosyntactic structure.

In (253) below the target and the basis of comparison are reversed:

- (253) *Se taitaa kuitenkin usein olla sitä "helmien heittämistä sioille", josta Jeesus varoitti (Matt. 7:6). Tällä en halua haukkua esim. Sinua siaksi, vaan totean, että sika ymmärtää helmistä yhtä paljon kuin sydämensä totuudelta lukinnut ihminen Jumalan ihmeistä.*<sup>368</sup>

lit. 'Anyway, I think it is often a case of "throwing pearls to pigs" of which Jesus warned (Matt. 7:6). By this I don't want to call e.g. you a pig, but I am stating that a pig understands about pearls as much as a person who has locked his heart from the truth (understands) about God's wonders.'

Another construction which is able to host both lexical items (i.e. *helmi* 'pearl' and *hopealusikka* 'silver spoon') and which is more or less a straightforward morphosyntactic realization of the CS [NEG[UNDERSTAND([ARG],[ARG])]] is *SIKA/SIAT EI(VÄT) YMMÄRRÄ HELMIEN/ HOPEALUSIKAN PÄÄLLE* | *NP<sub>SUBJ</sub>[N{sika 'pig'}] NEG{ei 'not'} Vcogn{ymmärtää 'understand'} PP<sub>PÄÄLLE</sub>[NP<sub>GEN SG/PL</sub>{helmi 'pearl'/ hopealusikka 'silver spoon'}]* | lit. 'pig(s) do(es) not understand about pearls/a silver spoon' as shown in Examples (254) and (255) below:

- (254) *Ei siat ymmärrä helmien päälle :-)*<sup>369</sup>  
lit. 'Pigs do not understand about pearls'  
id. 'A primitive, ignorant person does not appreciate something fine'
- (255) *Ei sika hopealusikan päälle mitään ymmärrä*<sup>370</sup>  
lit. 'A pig does not understand anything about a silver spoon'  
id. 'A primitive, ignorant person does not appreciate something fine'
- (256) *Toisaalta... jos kirjoitan tänne, että "sioille ei kannata helmiä heitellä", joku voi ymmärtää sen niin, että nimittelen jotakuta sikamaiseksi, alhaiseksi, sontaiseksi olennoiksi - vaikka siitä ei sillä kertaa olisikaan kysymys. Voihan olla, että näen toisen mukavan vaaleanpunaisena, kärsällään nuuhkivana, ystävällisenä ja sellaisena, jonka korvantaustaa tekee mieli rapsuttaa, mutta samalla tiedän, ettei hän helmien päälle mitään "ymmärtäisi" - eikä tarvoitsisikaan ymmärtää.*<sup>371</sup>

<sup>368</sup> GG: sfnet.keskustelu.uskonto.kristinusko/msg/6c53630c8c42845b, Aug 24 2000

<sup>369</sup> GG: sfnet.harrastus.elektroniikka/msg/a9b2a6506901405e, Jun 11 2003

<sup>370</sup> Sananparsikokoelma (2010) - Hausjärvi, J. Sillanpää, 1936-37

<sup>371</sup> GG: sfnet.keskustelu.psykologia/msg/f10cd6be89e28f00, Mar 6 1999

lit. 'On the other hand... if I write here, that **"it's no use throwing pearls to pigs"**, somebody can understand it to mean that I am calling somebody a piggish, low, mucky creature – even though that would not be the case. It may be that I see them as nice and pink, sniffing with their snout, friendly and making me feel like scratching behind their ear, but at the same time I know that **they would neither "understand" anything about pearls** – nor would they have to understand.'

- (257) *Usein autolehdissä moititaan XJ-mallien tilankäyttöä ja siihen ei voi kuin todeta: eipä se sika ole helmien päälle enenkään ymmärtänyt.*<sup>372</sup>

lit. 'In car magazines they often criticize the space utilization of XJ-models and to that one cannot help but note: **the pig has never understood much about pearls.**'

- (258) *Tietenkin sillä \_itselle\_ aina on, mutta jos sika ei helmistä mitään ymmärrä, eikä hyvä- ja huonolaatuisessa lopputuloksessa eroa näe, niin minkäs teet. Tuskin tilanne tällöin paranee teoriaa selittämälläkään.*<sup>373</sup>

lit. 'Of course, it always does (matter) \_for oneself\_, but if **a pig does not understand anything about pearls**, nor see the difference between a high- and the low-quality result, what can you do. In this case the situation can hardly be improved by explaining the theory.'

- (259) *Turha on heittää helmiä sioille. Ne ei helmien päälle ymmärrä - luulevat niitä vain huonosti sulavaksi ruuaksi.*<sup>374</sup>

lit. 'It's useless to **throw pearls to pigs. They do not understand about pearls** – (they) just think these to be poorly digestible food.'

The following Examples (260) – (262) demonstrate that *HELMIÄ SIOILLE* constructions are also able to host the interaction, i.e. to provide with the morphosyntactic structure, while the other construction family functions as a donor that contributes with the lexical item *hopealusikka* 'silver spoon':

- (260) *Enpä taida enään herra Burmaniin soveltaa ironiaa, sehän on kuin hopealusikoita sioille.*<sup>375</sup>

lit. 'I am not likely to apply irony to Mr. Burman, **that is like silver spoons to pigs.**'

- (261) *Mieheni tuo minulle kaakaota tähän, joten en kirjoittele enempää, koska on turha heittää hopealusikoita sioille, sillä kiribati, joka on kaikkien pissisten*

<sup>372</sup> GG: sfnet.harrastus.autot/msg/18dd38c87dd9a19b, Dec 13 2002

<sup>373</sup> GG: sfnet.atk.grafiikka/msg/3e60d3e2bb1efbbd, Sep 7 2006

<sup>374</sup> GG: sfnet.keskustelu.yhteiskunta/msg/b4960fff81b41db7, Nov 1 2006

<sup>375</sup> GG: sfnet.atk.sodat/msg/4d5ef2ebb5e7a858, May 22 1997

*ylipapitar, on täällä messunnut koko päivän, kun minä olen perheeni kanssa tehnyt työtä aamusta iltaan.*<sup>376</sup>

lit. 'My husband is bringing me hot chocolate here, so I will not write anymore, because **it is pointless to throw silver spoons to pigs**, as Kiribati, who is the high priestess of all chavs, has been here chanting the whole day long, while my family and I have been doing work from dawn to dusk.'

(262) A: *You're posting hard, verifiable evidence in response to a MJF posting. Have you ever heard the English saying "pearls before swine"?*<sup>377</sup>

B: *The Finnish one is "silver before swine". But nowadays almost everyone has adapted the Anglism and talks about pearls. Damn cultural imperialism.*<sup>378</sup>

C: *Really? Never heard of that silver version before... Kind of makes sense, since pearls were \_very\_ rare and hard to get here, but just how far back this idiom goes? Is it originally an adaption from Swedish, perhaps?*<sup>379</sup>

D: *Prolly from the Swedes. Actually, it's silver \*spoons\*... I don't think folks used that much silver cutlery before the middle age and Swedish-christian aggression anyway. Except for the worlords and kings and such who traded with southerners and easterners. "Turha syöttää/heittää sikaa hopealusikalla"<sup>a</sup> or just "Hopealusikoita sioille"<sup>b</sup> ie. "Don't bother feeding swine with a silver spoon" or just "Silver spoons before swine".<sup>380</sup>*

Three of the above examples feature two *HHS* constructions, where a syntactic constituent of the category noun, which by default corresponds to the lexical item *helmi* 'pearl', is now linked to a non-default donor lexical item *hopealusikka* 'silver spoon'. These constructions are:

- a. <W ON KUIN> HELMIÄ SIOILLE | <NP<sub>SUBJ</sub>{W} COP V{olla 'be'} COMP{kuin 'like'}> NP<sub>PTV</sub>[NP<sub>PL</sub>{helmi 'pearl'}] [PP<sub>ALL</sub>[NP<sub>PL</sub>{sika 'pig'}]] | lit. '<W is like> pearls to pigs' in Examples (260) and (262)-D-b;
- b. X HEITTÄÄ HELMIÄ SIOILLE | NP<sub>SUBJ</sub>[N{X}] V{heittää 'throw'} NP<sub>OBJ</sub>[NP<sub>PL</sub>{helmi 'pearl'}] PP<sub>ALL</sub>[NP<sub>PL</sub>{sika 'pig'}] | lit. 'X throws pearls to pigs' in Example (261).

In Example (262)-D<sup>a</sup> above, the host is not a *HHS* construction, but rather another Finnish PU – X HEITTÄÄ Y:LLÄ VESILINTUA | NP<sub>SUBJ</sub>[N{X}] V{heittää 'throw'} PP<sub>ADE</sub>[NP{Y}] NP<sub>OBJ PTV</sub>[N<sub>SG</sub>{vesilintu 'waterfowl'}] | lit. 'X throw Y at a waterfowl', id. 'X gets rid of Y (due to its uselessness)'. The similarity between this construction and X HEITTÄÄ HELMIÄ SIOILLE | NP<sub>SUBJ</sub>[N{X}] V{heittää 'throw'} NP<sub>OBJ</sub>[NP<sub>PL</sub>{helmi 'pearl'}] PP<sub>ALL</sub>[NP<sub>PL</sub>{sika 'pig'}] | lit. 'X

<sup>376</sup> <http://keskustelu.suomi24.fi/node/5715169#comment-30347019> [Accessed on 2.10.2008]

<sup>377</sup> GG: rec.autos.sport.f1/msg/b08a8c90a555399a, Aug 13 2000

<sup>378</sup> GG: rec.autos.sport.f1/msg/e57c24812802f240, Aug 13 2000

<sup>379</sup> GG: rec.autos.sport.f1/msg/cdf29631a4df4f20, Aug 13 2000

<sup>380</sup> GG: rec.autos.sport.f1/msg/aef3ad11d459cde7, Aug 14 2000

throws pearls to pigs’ is partially syntactic (both are SVO structures) and partially lexical (constructions share the lexical item *heittää* ‘throw’).

Figure 128 and Figure 129 below present a formal description of the interacting constructions (PU<sub>I</sub> and PU<sub>II</sub>) and resulting units (PU<sub>RES</sub>). In Figure 128 the host is Y EI YMMÄRRÄ/TIEDÄ Z:STA ENEMPÄÄ KUIN SIKA HOPEALUSIKASTA | NP<sub>SUBJ</sub>[N{Y}] NEG{ei ‘not’} Vcogn{ymmärtää ‘understand’/tietää ‘know’} PP<sub>ELA</sub>[NP{Z}] COMP{enempää kuin ‘more than’} NP<sub>SUBJ</sub>{sika ‘pig’} PP<sub>ELA</sub>[NP{hopealusikka ‘silver spoon’}] | lit. ‘Y does not understand Z/know about Z more than a pig about a silver spoon’, id. ‘Y does not understand Z/know about Z at all’ and the donor is X HEITTÄÄ HELMIÄ SIOILLE | NP<sub>SUBJ</sub>[N{X}] V{heittää ‘throw’} NP<sub>OBJ</sub>[N<sub>PL</sub>{helmi ‘pearl’}] PP<sub>ALL</sub>[NP<sub>PL</sub>{sika ‘pig’}] | lit. ‘X throws pearls to pigs’, id. ‘X causes a transfer of some entity, evaluated by the speaker as good, to a recipient, evaluated as inadequate in some way’. In Figure 129 the situation is reversed – the host is X HEITTÄÄ HELMIÄ SIOILLE | NP<sub>SUBJ</sub>[N{X}] V{heittää ‘throw’} NP<sub>OBJ</sub>[N<sub>PL</sub>{helmi ‘pearl’}] PP<sub>ALL</sub>[NP<sub>PL</sub>{sika ‘pig’}] | lit. ‘X throws pearls to pigs’, id. ‘X causes a transfer of some entity, evaluated by the speaker as good, to a recipient, evaluated as inadequate in some way’ and the donor is Y EI YMMÄRRÄ/TIEDÄ Z:STA ENEMPÄÄ KUIN SIKA HOPEALUSIKASTA | NP<sub>SUBJ</sub>[N{Y}] NEG{ei ‘not’} Vcogn{ymmärtää ‘understand’/tietää ‘know’} PP<sub>ELA</sub>[NP{Z}] COMP{enempää kuin ‘more than’} NP<sub>SUBJ</sub>{sika ‘pig’} PP<sub>ELA</sub>[NP{hopealusikka ‘silver spoon’}] | lit. ‘Y does not understand Z/ know about Z more than a pig about a silver spoon’, id. ‘Y does not understand Z/know about Z at all’. Both figures demonstrate that a resulting construction (PU<sub>RES</sub>) inherits the morphosyntactic structure of the host, as well as its conceptual structure.

The rule-based CS of KSH constructions is formalized as a relation between two propositions – UNDERSTAND([ARG][ARG]) and UNDERSTAND([PIG][PEARLS]) connected by a relational operator ≤ meaning ‘not more than’. This is a temporary technical solution. The whole problem of logical relations between two or more conceptual clauses within the same conceptual structure is a separate object of study, which deserves more attention. However, I am not able to dwell upon it in the scope of this book.

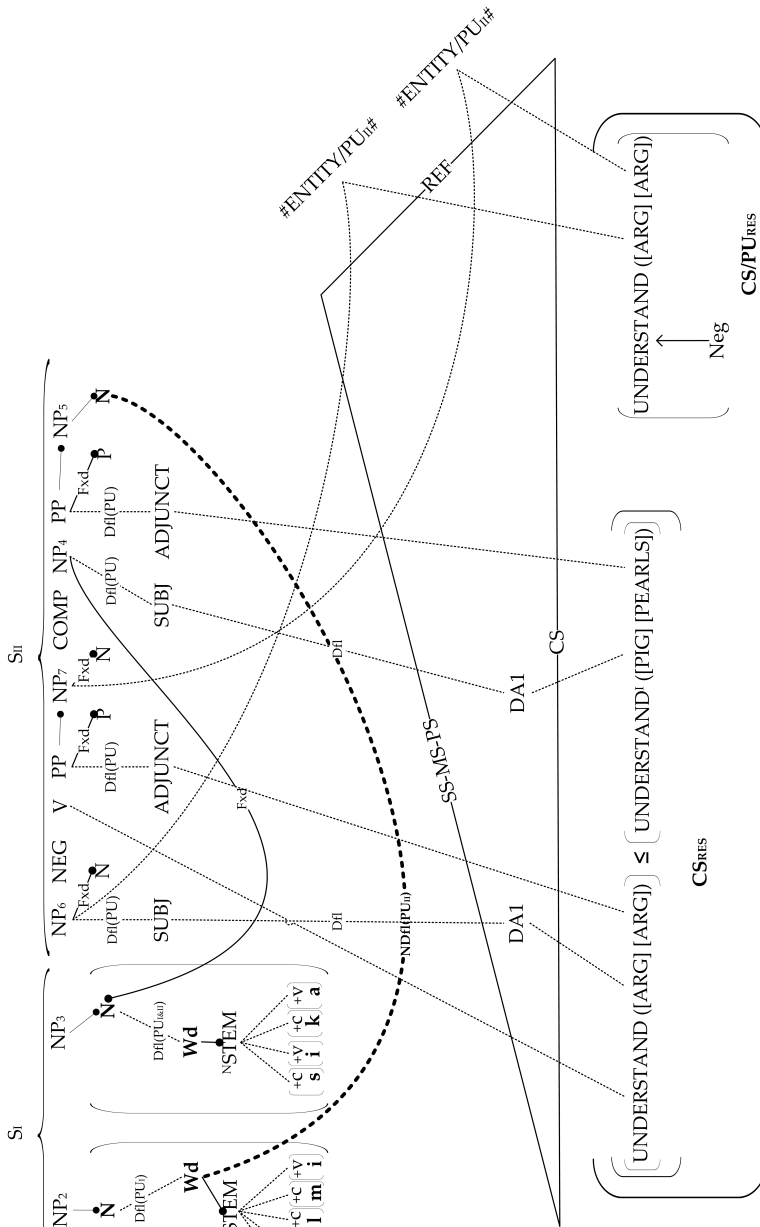


Figure 128 HOST = Y EI YMMÄRRÄ/TIEDÄ Z:STA ENEMPÄÄ KUIN SIKA HOPEALUSIKASTA lit. 'Y does not understand Z/know about Z more than a pig about a silver spoon', DONOR = X HEITTÄÄ HELMIÄ SIOILLE lit. 'X throws pearls to pigs'

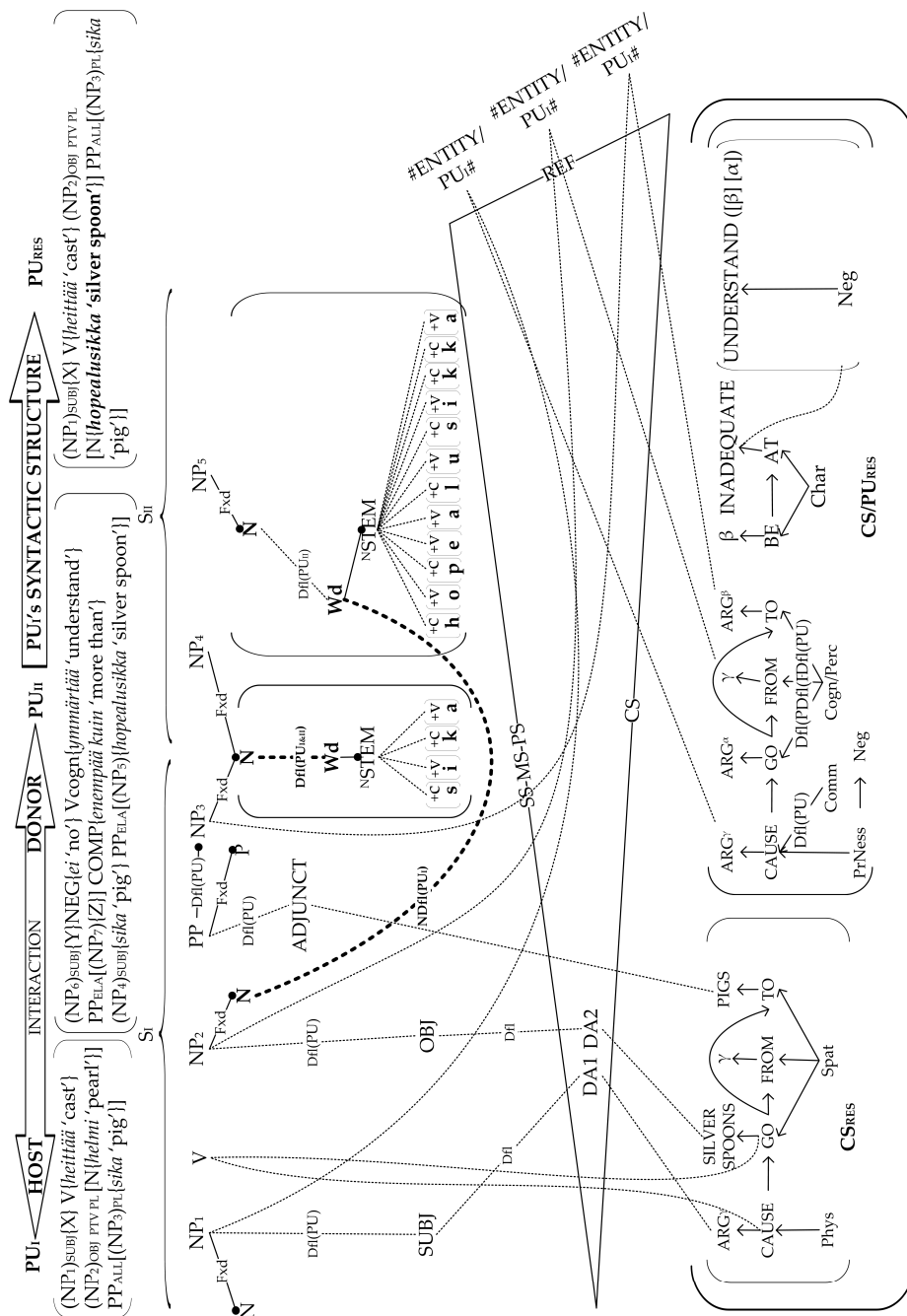


Figure 129 **HOST** = *X HEITTÄÄ HELMIÄ SIOILLE* lit. 'X throws pearls to pigs', **DONOR** = *Y EI YMMÄRRÄ/TIEDÄ Z:STA ENEMPÄÄ KUIN SIKÄ HOPEALUSIKASTA* lit. 'Y does not understand Z/know about Z more than a pig about a silver spoon'

There are several other constructions, which can interact as donors with KSH construction family. The results of such interaction are presented in Examples (263) – (265) below and summarized in Figure 130.

- (263) *Poliisitoimi perustuu aina ja kaikkialla oikeudenmukaisuuteen ja tasapuolisuuteen, josta sinä tiedät saman verran kuin sika jäniksestä.*<sup>381</sup>  
lit. ‘Police action is always and everywhere based on fairness and equality, of which you **know as much as a pig** about a rabbit.’
- (264) *Havaitsemme että tiedät ihmismielen pimeästä puolesta yhtä vähän kuin kala polkupyörästä.*<sup>382</sup>  
lit. ‘We see that you **know about** the dark side of the human mind **as little as a fish** about a bicycle.’
- (265) *Se joka sortuu niistä maksamaan ei varmaan tiedä asioista sen enempää kuin pässi uudesta veräjästä.*<sup>383</sup>  
lit. ‘Whoever resorts to paying for these probably **does not know any more about** things **than a ram** about a new gate.’

Here lexical correspondences between the noun constituent in  $PP_{ELA}[NP_1]$  and *polkupyörä* ‘bicycle’, *uusi veräjä* ‘new gate’ and *jänis* ‘hare’ are inherited from the following donor constructions:

- a. *SAA NÄHDÄ/KUIULLA SAAKO SIKÄ JÄNISTÄ* |  $V_{PRES\ ACT\ SG3}\{saada\ ‘get’\}$   $V_{INF}\{nähdä\ ‘see’/kuulla\ ‘hear’\}$   $V_{PRES\ ACT\ SG3\ KO}\{saada\ ‘get’\}$   $NP_{SUBJ}\{sika\ ‘pig’\}$   $NP_{OBJ\ PTV}\{jänis\ ‘hare’\}$  | lit. ‘we’ll see whether the pig will get the hare’, id. ‘it is doubtful, questionable’ (Example (263) above)
- b. *X TARVITSEE Y:TÄ YHTÄ PALJON/VÄHÄN KUIN KALA POLKUPYÖRÄÄ* |  $NP_{SUBJ}[N\{X\}]$   $V\{tarvita\ ‘need’\}$   $NP_{OBJ\ PTV}\{Y\}$   $COMP\{yhtä paljon/vähän kuin\ ‘as much/little as’\}$   $NP_{SUBJ}\{kala\ ‘fish’\}$   $NP_{OBJ\ PTV}\{polkupyörä\ ‘bicycle’\}$  | lit. ‘X needs Y as {much/little} as a fish (needs) a bicycle’, id. ‘X does not need Y’ (Example (264) above)
- c. *X KATSOO Y:TÄ KUIN PÄSSI/LEHMÄ UUTTA VERÄJÄÄ* |  $NP_{SUBJ}[N\{X\}]$   $V\{katsoa\ ‘look’\}$   $NP_{OBJ\ PTV}\{Y\}$   $COMP\{kuin\ ‘like’\}$   $NP_{SUBJ}\{pässi\ ‘ram’/lehmä\ ‘cow’\}$   $NP_{OBJ\ PTV}\{AP\{uusi\ ‘new’\} N\{veräjä\ ‘gate’\}\}$  | lit. ‘X looks at Y like a {ram/cow} at a new gate’, id. ‘X does not understand Y’ (Example (265) above).

<sup>381</sup> GG: sfnet.ryhmat+listat/msg/ce5a59d89e5420de, Oct 18 2001

<sup>382</sup> GG: sfnet.keskustelu.politiikka/msg/e5666e19d295d897, Jun 17 1998

<sup>383</sup> <http://keskustelu.suomi24.fi/node/2184365#comment-11028508>, Oct 19 2005





Of these three donor constructions only *SAA NÄHDÄ/KUULLA SAAKO SIKÄ JÄNISTÄ* | *V<sub>PRES ACT SG3</sub>{saada ‘get’} V<sub>INF</sub>{nähdä ‘see’/kuulla ‘hear’} V<sub>PRES ACT SG3 KO</sub>{saada ‘get’} NP<sub>SUBJ</sub>{sika ‘pig’} NP<sub>OBJ</sub> PTV{jänis ‘hare’}* | lit. ‘we’ll see whether the pig will get the hare’, id. ‘it is doubtful, questionable’ shares a lexical constituent *sika* ‘pig’ with the host construction. This lexical item provides practically the only explicit link between the two constructions, since their conceptual structures are not similar. The other two donor constructions do not share any lexical items with the host; however, the interaction is possible due to the fact that their conceptual structures involve elements similar to those of the host.

Examples (266) and (267) from Sananparsikokoelma (2010) show that the construction *X KATSOO Y:TÄ KUIN PÄSSI/LEHMÄ UUITTA VERÄJÄÄ* | *NP<sub>SUBJ</sub>[N{X}] V{katsoa ‘look’} NP<sub>OBJ</sub> PTV{Y} COMP{kuin ‘like’} NP<sub>SUBJ</sub>{pässi ‘ram’/ lehmä ‘cow’} NP<sub>OBJ</sub> PTV{AP{uusi ‘new’} N{veräjä ‘gate’}}* | lit. ‘X looks at Y like a {ram/cow} at a new gate’, id. ‘X does not understand Y’ is also able to host lexical items from the *KSH* construction family:

(266) *KATTELEE NINKUN SIKÄ TUULIMYLLYÄ*<sup>384</sup>  
lit. ‘looks like a pig at a windmill’

(267) *KATTOO KU SIKÄ HOPIA LUSIKKAA*<sup>385</sup>  
lit. ‘looks like a pig at a silver spoon’

The host-donor interaction can also result in an elliptic construction, as in the following example (268), where the Finnish *PU X OSTAA SIKÄ SÄKISSÄ* | *NP<sub>SUBJ</sub>[N{X}] V{ostaa ‘buy’} NP<sub>OBJ</sub>{sika ‘pig’} PP<sub>INE</sub>[NP{säkki ‘sack’}]* | lit. ‘X buys a pig in a sack’, id. ‘X buys something without inspecting the item beforehand and without knowing its true nature or value’ appears as a host, while the *HHS* construction family provides with a lexical item *helmi* ‘pearl’. This particular example features VP ellipsis in a negative polarity construction *ei ... vaan ...* ‘not ... but ...’, where the first negative part denies a proposition followed by a *vaan* conjunction and the second affirmative part, which presents a substitutive alternative to the proposition. The identical VP element is usually omitted in the substitutive construction (Hakulinen et.al 2004: 1138). In this case, the *NP<sub>OBJ</sub>{sika ‘pig’}* appears within the scope of negation in the first part, while in the substitutive construction the *NP<sub>OBJ</sub>{helmi ‘pearl’}* is presented

<sup>384</sup> Sananparsikokoelma (2010) - Tyrvää, F. Törmä, 1933

<sup>385</sup> Sananparsikokoelma (2010) - Rovaniemi, A. Alanampa, 1933

as an alternative. The  $PP_{INE}[NP\{säkki \text{ 'sack'}\}]$  is present only in the substitutive construction:

- (268) *Oliskohan ideaa jos joku pitäisi jonkinlaista ohjehinnastoa käytetyille synille? hmm.. toisaalta nyt säilyvä yllätysmomentti on kiva kun voi joskus saada ei sikaa, vaan helmen säkissä.. ;)*<sup>386</sup>

lit. 'Would it be a good idea if someone kept some kind of reference price list for used synths? hmm.. on the other hand, now the remaining moment of surprise is nice as you can sometimes **get not a pig, but a pearl in a poke.**'

The interaction in (268) above is possible due to the shared lexical constituent *sika* 'pig' and the juxtaposition between *helmi* 'pearl' as something valuable and *sika* 'pig' as something less valuable. Thus, the non-default lexical item *helmi* 'pearl' changes the default meaning of the host PU 'to make a risky purchase without inspecting the item beforehand' into 'to make a risky purchase without inspecting the item beforehand and unexpectedly get a valuable item'.

A similar negative polarity construction *ei ... vaan ...* 'not ... but ...' can be found in Example (269) below, where the second interacting construction is  $X \text{ MENEE } TAIVAAN \text{ TUULIIN } | NP_{SUBJ}[N\{X\}] \text{ } V\{\text{mennä 'go'}\} PP_{ILL}[NP_{PL}[NP_{GEN}\{taivas \text{ 'sky, heaven'}\} N\{tuuli \text{ 'wind'}\}]] |$  lit. 'X goes to heaven's winds', id. 'X disappears without a trace':

- (269) *vituttaa kaikkien niiden viestien poistaminen, joissa tidän olleen myös asiaa, jolloin helmiä ei ole mennyt edes sioille vaan taivaan tuuliin.*<sup>387</sup>

lit. 'I'm pissed off about the removing of all these messages that I know also contained relevant facts, meaning that **pearls haven't even gone to pigs, but to the four winds.**'

Here the elliptical substitutive construction, which appears after *vaan* 'but', contains only  $TAIVAAN \text{ TUULIIN } PP_{ILL} | [NP_{PL}[NP_{GEN}\{taivas \text{ 'sky, heaven'}\} N\{tuuli \text{ 'wind'}\}]] |$  lit. 'to heaven's winds'. An analogous elliptical  $PP_{ALL}$  can be observed in (270) below, the second interacting construction being the Finnish PU  $X \text{ MENEE } HARA KOILLE | NP_{SUBJ}[N\{X\}] \text{ } V\{\text{mennä 'go'}\} PP_{ALL}[NP_{PL}\{harakka \text{ 'magpie'}\}] |$  lit. 'X goes to magpies', id. 'X is wasted':

- (270) *Sinun viestisi noin yleisesti ottaen ovat todellisia helmiä, joita ei tässä tapauksessa heitetä sioille, vaan harakoille, sillä ei tässä ryhmässä ole muita kuin me.*<sup>388</sup>

<sup>386</sup> GG: sfnet.harrastus.musiikki.tekeminen/msg/f4233465287d3648, Oct 2 2001

<sup>387</sup> GG: sfnet.keskustelu.psykologia/msg/f0495d7d2b9d655c, May 31 2006

<sup>388</sup> GG: sfnet.harrastus.pelit.shakki/msg/97350a6c0a08edba, Jul 26 1998

lit. ‘Your messages generally speaking are real **pearls, which in this case are not thrown to pigs, but to magpies**, for in this group there is no one but us.’

The host-donor distinction, which has been presented in this section, is applicable to such cases, where one of the constructions (host) preserves its morphosyntactic structure, while the other is represented by a donated lexical item. However, there are cases where such distinction is not so straightforward (if at all possible), because both interacting constructions initially have similar morphosyntax. For instance, the morphosyntactic structure  $NP_{SUBJ} V\{mennä \text{ ‘go’}\} PP_{ALL}[NP_{PL}]$  is present in the following constructions:

- a. *HELMIÄ MENEE SIOILLE* |  $NP_{SUBJ} PTV[N_{PL}\{helmi \text{ ‘pearl’}\}] V\{mennä \text{ ‘go’}\} PP_{ALL}[NP_{PL}\{sika \text{ ‘pig’}\}]$  | lit. ‘pearls go to pigs’;
- b. *X MENEE HARAKOILLE* |  $NP_{SUBJ}[N\{X\}] V\{mennä \text{ ‘go’}\} PP_{ALL}[NP_{PL}\{harakka \text{ ‘magpie’}\}]$  | lit. ‘X goes to magpies’, id. ‘X is wasted’.

In Example (271) below the constituent  $NP_{SUBJ} PTV[N_{PL}\{helmi \text{ ‘pearl’}\}]$  appears together with the  $PP_{ALL}[NP_{PL}\{harakka \text{ ‘magpie’}\}]$  within the above-mentioned structure:

- (271) *Pilkkusääntöjä ei kuitenkaan kannatane YLEn radio- eikä tv-toimittajille opettaa - menee helmiä harakoille -- mutta painetun sanan toimittajille kyllä:-)*<sup>389</sup>  
 lit. ‘It’s pointless to teach comma rules to YLE’s radio and television journalists – **pearls go to magpies** – but to journalists of the printed word it is not:-)’

The structural difference between the two constructions lies in the fact that the morphological structure of the second construction is less specific: *X MENEE HARAKOILLE* |  $NP_{SUBJ}[N\{X\}] V\{mennä \text{ ‘go’}\} PP_{ALL}[NP_{PL}\{harakka \text{ ‘magpie’}\}]$  | lit. ‘X goes to magpies’, id. ‘X is wasted’ does not specify any particular default morphological form for its  $NP_{SUBJ} PTV[N_{PL}\{helmi \text{ ‘pearl’}\}]$  in the resulting structure is inherited from *HELMIÄ MENEE SIOILLE* |  $NP_{SUBJ} PTV[N_{PL}\{helmi \text{ ‘pearl’}\}] V\{mennä \text{ ‘go’}\} PP_{ALL}[NP_{PL}\{sika \text{ ‘pig’}\}]$  | lit. ‘pearls go to pigs’. Therefore, since the resulting structure has more morphosyntactic features of the latter construction, it could technically be regarded as a host.

In Example (272), *PESUVEDEN SEASSA* |  $PP_{SEASSA}[NP_{GEN}\{pesuvesi \text{ ‘bathing water’}\}]$  | lit. ‘among bathing water’, which appears together with the

<sup>389</sup> GG: sfnet.ryhmat+listat/msg/1928345eb46c1cfa, Apr 29 2002

$NP_{SUBJ\ PTV}[N_{PL}\{\textit{helmi}\ 'pearl'}\}]$ , originates from the Finnish PU  $LAPSI\ MENE\ PESUVEDEN\ SEASSA$  |  $NP_{SUBJ\ NOM}[N_{SG}\{\textit{lapsi}\ 'child'}\}] V\{\textit{mennä}\ 'go'}\} PP_{GEN\ SEASSA}[NP_{SG}\{\textit{pesuvesi}\ 'bathing\ water'}\}]$  | lit. 'a child goes out with the bathing water', id. 'some good parts become wasted, lost when one gets rid of the bad parts of something':

(272) *Ajoin asiaa hyvin tietoisena siitä, että siinä varmasti menee helmiä pesuveden seassa [...]*<sup>390</sup>

lit. 'I fought the case knowing well that **pearls will certainly go out with the bathing water [...]**'

Morphosyntactic structures of these interacting constructions differ in the default forms of their  $NP_{SUBJ}$  constituents ( $NP_{PTV}[N_{PL}]$  vs.  $NP_{NOM}[N_{SG}]$ ) and PP constituents ( $PP_{ALL}[NP_{PL}]$  vs.  $PP_{GEN\ SEASSA}[NP_{SG}]$ ). The resulting structure in Example (272) above inherits both phonological and morphological specifications of the  $NP_{SUBJ}$  from  $HELMIÄ\ MENE\ SIOILLE$  |  $NP_{SUBJ\ PTV}[N_{PL}\{\textit{helmi}\ 'pearl'}\}] V\{\textit{mennä}\ 'go'}\} PP_{ALL}[NP_{PL}\{\textit{sika}\ 'pig'}\}]$  | lit. 'pearls go to pigs' construction, while the PP constituent in its phonological and morphological form is inherited from  $LAPSI\ MENE\ PESUVEDEN\ SEASSA$  |  $NP_{SUBJ\ NOM}[N_{SG}\{\textit{lapsi}\ 'child'}\}] V\{\textit{mennä}\ 'go'}\} PP_{GEN\ SEASSA}[NP_{SG}\{\textit{pesuvesi}\ 'bathing\ water'}\}]$  | lit. 'a child goes out with the bathing water' construction. Thus, technically no host-donor distinction can be made, since both interacting constructions are equally represented in the resulting structure.

Host-donor interaction in (272) above and in (273), (274) and (275) below is based on the conceptual contraposition between something important, essential and something unimportant, inessential, which is present in the CS/PUs of interacting constructions. In Example (273) below the donor construction is the Finnish PU  $KYLLÄ\ SOKEAKIN\ KANA\ JOSKUS\ JYVÄN\ LÖYTÄÄ$  lit. 'Even a blind chicken sometimes finds a grain', id. 'even a less worthy may succeed, may get a good idea etc.'. The host here is the intransitive HHS construction  $HELMIÄ\ MENE\ SIOILLE$  |  $NP_{SUBJ\ PTV}[N_{PL}\{\textit{helmi}\ 'pearl'}\}] V\{\textit{mennä}\ 'go'}\} PP_{ALL}[NP_{PL}\{\textit{sika}\ 'pig'}\}]$  | lit. 'pearls go to pigs'. Lexical items *sokea* 'blind', *kana* 'chicken' and *löytää* 'find' borrowed from the donor constructions appear in the resulting constructions as an agent participle attribute of the  $SOKEAN\ KANAN\ LÖYTÄMÄ\ HELMENJYVÄNEN$  |  $NP_{SUBJ}[PartisP[NP_{GEN}\ [AP\{\textit{sokea}\ 'blind'}\} N_{SG}\{\textit{kana}\ 'chicken'}\}] PartisMA\{\textit{löytää}\ 'find'}\}] N\{\textit{helmenjyvänen}\ 'pearl_{GEN}grain_{NEN},\ grain\ of\ pearl'}\}]$  | lit. 'a grain of

<sup>390</sup> GG: sfnet.harrastus.radio.ham/msg/72501c950f6e67f0, Aug 1 2002

pearl found by a blind chicken’. *Helmenjyvänen* ‘grain of pearl’ is a compound word which contains morphological stems of both *helmi* ‘pearl’ and *jyvä* ‘grain’:

- (273) *En kiellä kuitenkaan, etteikö sokean kanan löytämä helmenjyvänen olis mennyt tällä kertaa sioille.*<sup>391</sup>

lit. ‘However, I am not denying that a grain of pearl found by a blind chicken went to pigs this time’.

In Example (274) below the host is the Finnish PU *EI KANNATA TUHLATA/HAASKATA RUUTIA VARIKSIIN* | *NEG{ei ‘not’} MV{kannattaa ‘be worthwhile, behoove’} V{tuhlata ‘waste’/ haaskata ‘waste’} NP<sub>OBJ</sub> PTV[N<sub>SG</sub>{ruuti ‘gunpowder’}] PP<sub>ILL</sub>[NP<sub>PL</sub>{varis ‘crow’}]* | lit. ‘it’s no use wasting gunpowder on crows’:

- (274) *Ei sfnetin sikoihin helmiä haaskata, ainoastaan lokaa.*<sup>392</sup>

lit. ‘Let’s not waste pearls on sfnet’s pigs, only mud.’

In (275) below the host construction is the Finnish PU *X EI NÄE METSÄÄ PUULTA* | *NP<sub>SUBJ</sub>[N{X}] NEG{ei ‘not’} V{nähdä ‘see’} NP<sub>OBJ</sub> PTV[N<sub>SG</sub>{metsä ‘forrest’}] PP<sub>ABL</sub>[NP<sub>PL</sub>{puu ‘tree’}]* | lit. ‘X does not see the wood for the trees’ id. ‘X does not notice the main thing, the whole, the essential because of minutiae, details, non-essentials’ and the donor is the *HHS* construction family. Originally the interacting constructions do not share any lexical items, although some parallels can be found in their syntactic structure:

- (275) *A: Ei sitten tule bongattua niitä helmiäkään.*

*B: Ongelma lienee se, että täällä ei vain näe helmiä sioilta.*<sup>393</sup>

lit. ‘A: Then those pearls won’t be spotted either.

B: It seems that the problem is that one just cannot see pearls for the pigs here.’

### 5.3.4.2 Embedding and coordination

Example (276) below presents a case of interaction between two different constructions, somewhat different from those I have showed so far. Here one is no longer dealing with a single lexical item from the donor construction borrowed into a host morphosyntactic pattern, but instead a whole construction (in this case the *HHS* construction *HELMIÄ SIOILLE* | *NP<sub>PTV PL</sub>[N{helmi ‘pearl’}] PP<sub>ALL</sub>[NP<sub>PL</sub>[N{sika ‘pig’}]]]* | lit. ‘pearls to pigs’) is

<sup>391</sup> GG: sfnet.harrastus.elokuvat/msg/03f6f6eea994654dd, Nov 20 1996

<sup>392</sup> GG: sfnet.huuhaa/msg/b938528d10680a12, Sep 13 2005

<sup>393</sup> GG: sfnet.huuhaa/msg/bc69d7bfe6809f20, Oct 23 2003

embedded into a larger construction – a wellerism structure *X*, *SANOI Y*, *KUN Z* ‘*X*, said *Y*, when *Z*’ (see e.g. Östman 2002, who analyzes *Sulva* wellerisms as constructions). I call this kind of interaction embedding:

(276) “*Helmiä sioille*”, *sanoi insinööri, kun newsseihin kirjoitti*.<sup>394</sup>

lit. “‘**Pearls to pigs**”, said the engineer, when he wrote to the news(groups).’

The following examples involve interaction of two *HHS* constructions – *HELMIÄ SIOILLE* | *NP<sub>PTV PL</sub>[N{helmi ‘pearl’}] PP<sub>ALL</sub>[NP<sub>PL</sub>[N{sika ‘pig’}]]]* | lit. ‘pearls to pigs’ in (277) and *<NP<sub>SUBJ</sub>> Vtr{heittää ‘throw’} NP<sub>OBJ PTV PL</sub> [N{helmi ‘pearl’}] PP<sub>ALL</sub>[NP<sub>PL</sub> [N{sika ‘pig’}]]]* in (278), with two other Finnish PUs – *EI KANNETTU VESI KAIVOSSA PYSY* lit. ‘carried water does not stay in the well’, id. ‘knowledge which is received from others and not absorbed is easily forgotten’ represented by the deverbal NP construction *VEDEN KANTAMINEN KAIVON* | *NP[NP<sub>GEN SG</sub>[N{vesi ‘water’}]] DVtrN{kantaminen ‘carrying’} PP<sub>ILL</sub>[NP<sub>SG</sub>[N{kaivo ‘well’}]]]* | lit. ‘carrying water into a well’ in (277) and *KUIN KAATAISI VETTÄ HANHEN SELKÄÄN* lit. ‘as if (one) would pour water on a goose’s back’, id. ‘a vain, ineffective attempt to influence somebody, give an advice or criticize’ represented by the elliptical construction *VETTÄ HANHEN SELKÄÄN* | *NP<sub>OBJ PTV SG</sub>[N{vesi ‘water’}] PP<sub>ILL</sub>[NP[NP<sub>GEN</sub>[N{hanhi ‘goose’}] N{selkä ‘back’}]]]* | lit. ‘water on a goose’s back’ in (278). The interaction is not happening in a way I have observed in the previous section. Here none of the constructions is the host structure or is embedded into the other, and both preserve their default lexical constituents. At the same time, they are joined by a coordinating conjunction into a coordinative conjunction phrase (Korhonen 1993). The interaction is licensed by the shared modal *PrNess→Neg* features (Sections 3.4.3.6.2, 5.3.2) in CS/PU of these constructions.

(277) *Tai turhaa tämä selittäminen täällä on. Helmiä sioille ja veden kantamista kaivoon*.<sup>395</sup>

lit. ‘Or this explaining is pointless here. **Pearls to pigs and carrying water into a well.**’

(278) *Tieteen opettaminen Juha-Pekalle on kuin heittäisi helmiä sioille, tai vettä hanhen selkään*.<sup>396</sup>

<sup>394</sup> GG: sfnet.keskustelu.kuluttaja/msg/320d235cfc595988, Dec 19 1997

<sup>395</sup> GG: sfnet.keskustelu.liikenne/msg/b9107f78a16e1009, Jan 29 2002

<sup>396</sup> GG: sfnet.keskustelu.evolutio/msg/820f620acb00673c, Sep 14 2006

lit. 'Teaching science to Juha-Pekka is like **throwing pearls to pigs, or water on a goose's back.**'

## 5.4 Discourse-related variation in PU

In Section 5.2 I touched upon the notion of phraseological cohesion as a complex interrelation between a PU and its context. In a nutshell: PUs create cohesion in texts, in which they are used; at the same time discourse context coheres with a PU in the sense that textual elements correspond to parts of the PU's structure. In Section 3.4.3.4 of Chapter 3 I discussed some principles of formalization of the referential claims made by a PU. Following Jackendoff (2002), referential relations were analyzed by incorporating a referential tier into a PU's network structure. It turns out that this part of the structure can be effectively applied to the analysis of PU variation as well. Discourse context reference can have an impact on PU use in several ways:

- By explicating the contextual meaning of ambiguous PUs, such as *HHS* constructions. Section 5.4.1 below is dedicated to the analysis of referentially-licensed semantic variation in the CS/PU of *HHS* constructions (some preliminary analysis can be also found in Section 3.4.3.5 of Chapter 3).
- By directly affecting a PU's form, as happens with referentially licensed non-default constituents discussed in Section 5.4.2 below;
- By providing cohesive links to the PU in cases of extreme reduction of its form to single structurally unrelated constituents analyzed in Section 5.4.3 below.

### 5.4.1 #SITUATION#/PU and its linking to the CS/PU

In this Section I will look at CS/PU correspondences to the referential tier and examine the PU's ability to be used in different situations as well as some of the possible relations between the referent of *pearls* and the referent of *pigs*. Figure 91 in Section 5.1.3 gives some idea of variation in semantic fields of CS/PU and #ENTITY/PUs#. Being to a large extent unspecific and therefore having a "broad meaning potential" (Halliday 1985), the CS/PU of the construction family in question still imposes certain restrictions to the range of different situations which this unit can be used to conceptualize. The major restricting factors here are:



- The potential well-formed argument structure of a conceptualized situation largely follows the same logic as the argument structure of non-default verbs discussed in 5.3.3.3.2 of the present chapter.
- The potential conceptualized entities are well-formed if their referents can be evaluated in the same way as these entities are evaluated in the CS/PU (see Section 3.4.3.5 for preliminary discussion).

The following elements of the CS/PU are context-dependent and therefore can vary according to the conceptualized situation:

- Thematic arguments of the main conceptual clause;
- Semantic fields of the main conceptual clause;
- The property INADEQUATE assigned to the argument [ARG]<sup>β</sup> co-referential with  $NP_{PL}\{sika\ 'pig'\}$ .

Here we are dealing with explication of referential claims made about the conceptualized situation and its participants, which are otherwise unspecified. As has been pointed out in Chapter 3 Section 3.4.3.5, both the  $NP_{PL}\{helmi\ 'pearl'\}$  and  $NP_{PL}\{sika\ 'pig'\}$  are #SITUATION#/PU referential, i.e. they correspond on the referential tier to their own #ENTITY#/PU indices selected by the #SITUATION#/PU. Each #ENTITY#/PU index, in turn, corresponds in the CS/PU to an unspecified thematic argument [ARG]. These arguments are selected by the same thematic tier functions (GO and TO) as the conceptual arguments [PEARL] and [PIG] in the regular CS and thus occupy THEME and LANDMARK positions. Since none of the thematic arguments in the CS/PU is specified, the PU remains semantically ambiguous, unless the referential indeterminacies are resolved. The complete truth-conditional content in the form of a specific CS/PU can be achieved only if the respective #ENTITY#/PU indices corresponding to  $NP_{PL}\{helmi\ 'pearl'\}$  and  $NP_{PL}\{sika\ 'pig'\}$  are derived from the f-knowledge base, including the actual context in which the PU is used. Following Jackendoff (2002), I assume that the task of referential indices is to pick out the entities of the conceptualized world. Established referents provide semantic content to otherwise empty argument slots on the thematic argument tier of the CS/PU.

In what follows I will present several selected examples of different #SITUATION#/PU linkings grouped according to relations that exist between intended referents of *pearls* and intended referents of *pigs*. Although they can give some idea of the semantic diversity in this PU, this is by no means a complete picture. A more detailed analysis had to be left outside the scope of the present study.



lit. 'In their home markets, Japanese manufacturers do offer the most interesting vehicles, but apparently (they) have decided that **no pearls to pigs** and are exporting to Europe almost solely grey mass.'

- (281) A: *Ja koska Porsche on jättänyt itselleen oikeuden valita kyseisen mallin ostajat, tilaajan on pakko olla pidemmän linjan possumies (vaiko peräti -nainen näin tasa-arvon aikoina?).*<sup>399</sup>

B: *Tottakai. Miksi myydä helmiä sioille...vai pitäisikö sanoa possuille.*<sup>400</sup>

lit. 'A: And, since Porsche has reserved the right to choose buyers for the model in question, the orderer has to be an old-time piggy-man (or even a woman, in these times of equality?)

B: Sure. Why **sell pearls to pigs** ... or should I say to piggies.'

- (282) *helmi kun kuvaa oikein hyvin esim. uudenkarheaa Mersua, johon likainen ja siviistymätön maatalonisäntä sitten istahtaa raappahousuineen. Tuohon voi sitten autokauppias todeta, että "helmiä sioille" kun paiskaa käteisenä maksetun kauppahinnan kassaan.*<sup>401</sup>

lit. 'A pearl describes very well e.g. a brand new Benz, into which a dirty and uncivilized farm owner seats himself with his long johns. To that a car dealer can then say that "**pearls to pigs**" as he throws the purchase price paid in cash into the cash register.'

- (283) *Miten muuten voi olla mahdollista, että joku M3:en omistaja ei ole muka ajanut kuin 209 km/h? Helmiä sioille?*<sup>402</sup>

lit. 'By the way, how is it possible that some M3-owner supposedly has not been driving faster than 209 km/h? **Pearls to pigs?**'

2. #SITUATION#/PU = #SERVICE-PROVIDING#

#PEARL-THROWER# = #SERVICE PROVIDER#

#PEARL# = #SERVICE#

#PIG# = #CUSTOMER#

*PPO*, *Sonera* and *Elisa* appearing in (284) below are Finnish telecommunications SERVICE PROVIDERS. The referent of *pearls* is a faster broadband connection, i.e. a SERVICE and the referent of *pigs* are their CUSTOMERS. Since a SERVICE is the non-ownership equivalent of a PRODUCT there is no transfer of ownership and hence no possessive semantic field in the S-tier (Figure 132 below).

<sup>399</sup> GG: sfnet.harrastus.autot/msg/1bf3c08e2f82b6a8, Feb 28 2002

<sup>400</sup> GG: sfnet.harrastus.autot/msg/599b744db8a1390c, Feb 28 2002

<sup>401</sup> GG: sfnet.keskustelu.kieli.kaantaminen/msg/062f11d9dc433682, Dec 1 2000

<sup>402</sup> GG: sfnet.harrastus.autot/msg/67c85ef24a9e78b0, Oct 8 2002

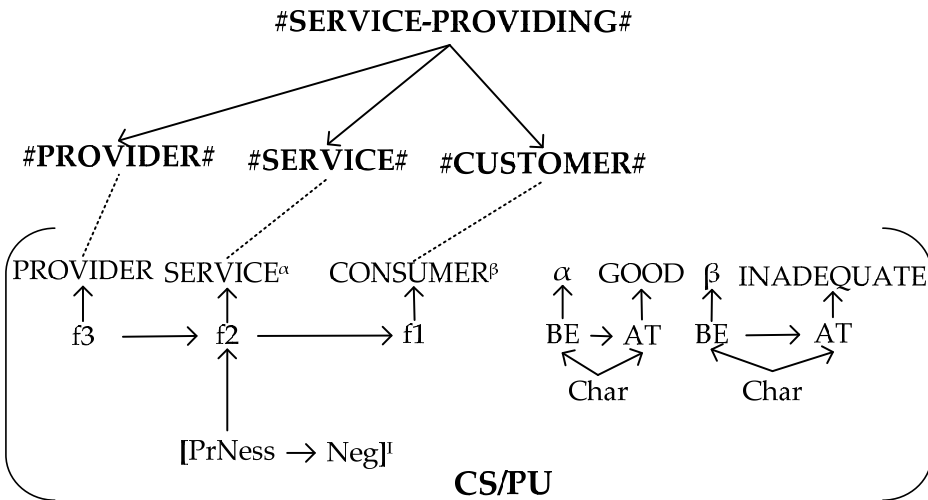


Figure 132 #SERVICE-PROVIDING#

(284) No, siinäpä se: täällä kun ei taida kukaan tarjota enempää kuin 4/jotakin, tämä PPO:lta. Ilmoitus taisi olla jossain ihan tuoreessa paikallislehdessä, ehkä Saarijärveläisessä? Minun käsittääkseni jopa tuo 4 megaa on PPO:lta ihan uutta, Soneralla on tarjolla kai edelleenkin enintään 2 megaa. Henkilökohtaisesti minä vähän epäilen että kilpailun puute on se Elisan "ongelma" tässä tapauksessa. :P Mitäs sitä suotta **tarjoamaan helmiä sioille**, kun ei kerran muutkaan.<sup>403</sup>

lit. 'Well, that's it: since no one here seems to offer more than 4/something, this is from PPO. I think the ad was somewhere in the latest local newspaper, maybe in Saarijärveläinen? As far as I know, even that 4 mega is quite new for PPO, I guess Sonera is still offering no more than 2 mega. Personally, I somewhat suspect that Elisa's "problem" in this case is the lack of competition :P Why bother to **offer pearls to pigs**, when no one else does.'

#### 5.4.1.2 Interpersonal relations

1. #SITUATION#/PU = #SOCIOEMOTIVE INTERACION#  
 #PEARL-THROWER# = #<INTIMATE> PARTNER#  
 #PEARL# = #SOCIOEMOTIVE RELATIONAL ELEMENTS#  
 #PIG# = #<INTIMATE> PARTNER#

The conceptualized situation in (285) below emphasizes a socioemotive interaction between partners. Socioemotive relational elements involve e.g. trust, respect, commitment, emotional intimacy etc.

<sup>403</sup> GG: sfnet.tietoliikenne.yhteydentarjoajat/msg/0946982e0e11df8b, Apr 7 2005

(285) A: *Olisiko niin, että jos jotain ihmistä kohtaan suhtaudutaan armottomasti/arvottomasti, hän alkaa toistamaan tätä suhtautumista myös kohtaamiinsa uusiin ihmisiin [...]?*

B: *On. En minä mikään ihmisvihaaja ole, mutta olen vuosien kuluessa tullut siihen tulokseen että minä en anna itsestäni yhtään enempää kuin on välttämätöntä koska menneinä vuosina **helmet menivät aina sioille** ja aikanaan väsyin siihen.*<sup>404</sup>

lit. 'A: Could it be that if some person is treated with ruthlessness/worthlessness, he begins to repeat this attitude also with new people that he encounters [...]?

B: Yes it could. I am no misanthrope, but over the years I have come to the conclusion that I do not give of myself any more than is necessary because in the past **pearls always went to pigs** and I eventually got tired of it.'

2. #SITUATION#/PU = #PARTNERSHIP#

#PEARL# = #<INTIMATE> PARTNER#

#PIG# = #<INTIMATE> PARTNER#

In (286) below the conceptualized situation involves both interpersonal relations and certain possessive attitudes/behaviour of one partner towards another partner in an intimate relationship. A pretty woman is treated as a valuable piece of private property, which has to be protected from encroachment:

(286) *Ei helmiä sioille. Eikä nättejä naisia sellaiselle, joka ei jaksata ottaa elämäntehtäväkseen hätistellä miljoonia himoissaan kuolaavia machomiehiä, joiden mielestä heidän kuuluu 'pelastaa' tuollaiset neidot omaan punkkaansa kutemaan.*<sup>405</sup>

lit. 'No pearls for pigs. Nor pretty women for someone who can't be bothered to make it their purpose in life to drive away millions of lustful drooling macho men, who believe that they have to 'save' such maidens into their own bunk to spawn.'

#### 5.4.1.3 Perception, experience

#SITUATION#/PU = #PERCEPTION/EXPERIENCE#

#PEARL# = #PHENOMENON#

#PIG# = #PERCEIVER/EXPERIENCER#

1. SENSE MODALITY = GUSTATORY& OLFACTORY

<sup>404</sup> GG: sfnet.keskustelu.psykologia/msg/ca214b8a082eca79, Jul 11 2001

<sup>405</sup> GG: sfnet.keskustelu.ihmissuhteet/msg/e78b202aceb571a0, Apr 16 2002

- (287) *Mutta minä en osaa ajatella aamukahvia muuna kuin osana aamiaista eli ateriala, joka nautitaan ennen heräämistä, joten makujen ja tuoksujen nyanssit ovat minun osaltani silloin lähinnä helmiä sioille.*<sup>406</sup>

lit. 'But I can't imagine morning coffee as anything other than a part of breakfast, or the meal which is eaten before waking up, so on my part nuances of flavours and aromas are then mostly **pearls to pigs**.'

## 2. SENSE MODALITY = AUDITORY

- (288) *Ollaan yritetty yhdessä puolustella klassisen musiikin arvokkuutta! Mä en enää jaksaa. Täysin käsittämätöntä, ettei ihmiset osaa arvostaa sitä mitenkään. kai tässä täytyy vain todeta että ei helmiä sioille vai mitä...*<sup>407</sup>

lit. 'We have tried together to defend the value of classical music! I can't go on anymore. It is totally incomprehensible that people cannot appreciate it in any way. I guess all that can be said is **no pearls to pigs**, or what...'

## 3. SENSE MODALITY = VISUAL

- (289) *A: Erehtynyt antamaan rahojaan tuonne - 8 euroa alkuperäisen pääsiäisaterian näkemisestä? Unohda koko homma.*<sup>408</sup>

*B: Aivan, taide ei ole kaikkia varten. Miten se menikään se juttu helmistä ja sioista...*<sup>409</sup>

lit. 'A: Made the mistake of giving his money there - 8 euro for seeing the original last supper? Forget the whole thing.'

B: Exactly, art is not for everyone. What was the story about **pearls and pigs**...'

### 5.4.1.4 Structure and function

#PEARL# = #{PART/FEATURE/FUNCTION}#

#PIG# = #{WHOLE/SYSTEM/STRUCTURE/USER}#

These situations mostly appear in the context of computer technology and gadgets. For example, in (290) below the referent of *pearls* is an installation program and the referent of *pigs* is *Debian* - a Linux operating system, which apparently does not work properly with this program:

- (290) *A: Asennusohjelmista puheenolleen, mikähän mahtaa olla ns. paras? InstallShieldin evaluointiversiota en Debianiini jostain kumman syystä saanut asennettua..*<sup>410</sup>

<sup>406</sup> GG: sfnet.harrastus.ruoka+juoma/msg/9dfb711a36c56e34, Sep 21 2004

<sup>407</sup> GG: sfnet.harrastus.musiikki.tekeminen/msg/2b6e0b3494a6fdb5, Aug 19 1999

<sup>408</sup> GG: sfnet.matkustaminen/msg/b7f5f6871939797a, May 13 2006

<sup>409</sup> GG: sfnet.matkustaminen/msg/8a7299d554921045, May 13 2006

<sup>410</sup> GG: sfnet.atk.sodat/msg/1c108803cddf4bba, Jul 19 2001

B: *ei helmiä sioille -;*)<sup>411</sup>

lit. 'A: Apropos installation programs, I wonder which one is the best? For some strange reason I was not able to install the evaluation copy of InstallShield on my Debian.

B: **no pearls to pigs -;**'

In (291) below the referent of *pigs* is the USER of a motorcycle, whose skills are not enough to exploit all potential features of the shock absorber:

(291) *Iskarin toiminta on juurikin sellaista kuin kuvittelinkin ja pelkäänpä, että minun taidoilla lisäsäädöt olisivat **kuin sioille heiteltyjä helmiä**.*<sup>412</sup>

lit. 'The shock absorber's functioning was just like I imagined and I am afraid that with my skills further adjustments would be like **pearl thrown to pigs**.'

#### 5.4.1.5 Verbal communication

#SITUATION#/PU = #COMMUNICATION#

#PEARL# = #SPEAKER#

#PIG# = #ADDRESSEE#

The semantic fields in the main conceptual clause of the CS/PU are Communicative and Cognitive-Perceptual, while the proposition itself encodes a communicative event.

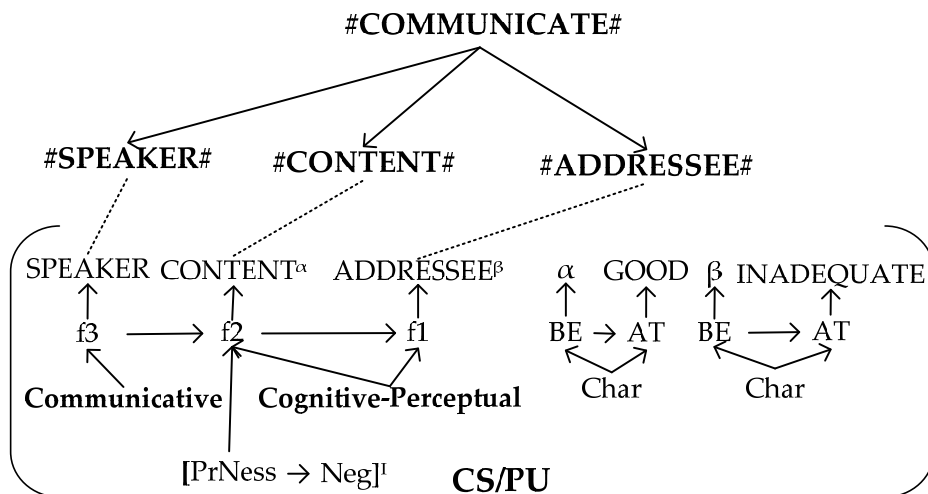


Figure 133 Communicative event

<sup>411</sup> GG: sfnet.atk.sodat/msg/3e82af1ce5069fd1, Jul 19 2001

<sup>412</sup> GG: sfnet.harrastus.mp/msg/68945348121f4b7a, Jan 4 2005

- (292) *Sinusta saa sellaisen kuvan, että olet tottunut paukuttamaan Ikuisia Totuuksia jostain hyvin korkealta armoitetulle yleisölle (opiskelijoillesi?), joka on niin taulapääistä, että perustelujen esittäminen on **helmien heittämistä sioille**.*<sup>413</sup>  
 lit. ‘You give the impression that you’re accustomed to cracking Eternal Truths from somewhere very high above to the blessed audience (your students?), which is so dumb that giving explanations would be **throwing pearls to pigs**.’

## 5.4.2 Non-default lexical linking and the referential tier

In PUs with #SITUATION#/PU referential constituents, such as the *HHS* construction family (see Sections 3.4.3.3 – 3.4.3.5 in Chapter 3 for preliminary discussion) there is a possibility for referentially licensed non-default correspondence that can be established between:

1. the PU’s noun constituent and phonological structure corresponding to the conceptual structure of a thematic argument in the CS/PU and corresponding to the intended #ENTITY#/PU (5.4.2.1 below), and
2. the PU’s verb constituent and phonological structure corresponding to the argument structure of the main conceptual clause in the CS/PU and corresponding to the intended #SITUATION#/PU (5.4.2.2 below).

Both types of non-default linking serve as means of disambiguation and explication of the CS/PU.

### 5.4.2.1 Referentially licensed non-default nouns

Consider the example in (293) below, where *NP<sub>PL</sub>* in *PP<sub>ALL</sub>* corresponds to the phonological form *Junnuille* ‘Junn<sub>ALL PL</sub>, to Junnus’ instead of the default *sioille* ‘pig<sub>ALL PL</sub>, to pigs’:

- (293) [*Minä*]<sub>1</sub> olen esittänyt pohdittavaksi, [*olisiko kansantalouden edun mukaista lisätä eläkerahastojen sijoituspolitiikkaan yksi pieni lisäriteeri, jolla kotimaahan tehtävän sijoituksen kannattavuutta arvioitaessa kotimaisen investoinnin kannattavuutta arvioitaisiin myös yhteiskuntahyötyjen näkökilmasta*]<sub>2</sub> - ja sinä olet kauhuissasi sosialismista, jota [*tuo*]<sub>2</sub> tarkoittaisi. Musta tuntuu, että se olen [*[mina]*]<sub>1</sub>, joka syyttää [*helmiä*]<sub>2</sub> [*Junnuille*]<sub>3</sub>]<sub>4</sub>, kun [*suostu[n]*]<sub>1</sub> [*sulle*]<sub>3</sub> [*tätä asiaa*]<sub>2</sub> edelleenkin selvittämään]<sub>4</sub><sup>414</sup>

lit. ‘[I]<sub>1</sub> have presented for consideration (a question) [whether it is in the interest of the national economy to add to the investment policy of pension funds one small additional criterion, according to which, when assessing the profitability of investment that has to be made in the home country,

<sup>413</sup> GG: sfnet.keskustelu/msg/48b1b4c93088278d, Feb 15 1996

<sup>414</sup> GG: sfnet.keskustelu.talous/msg/bd627ca514f9016f, Feb 28 2006



the profitability of domestic investment would be assessed also from the point of view of benefits to society]<sub>2</sub> - and you are terrified by socialism, which [that]<sub>2</sub> stands for. I feel that it's **[[me]<sub>1</sub>, who ladles out [pearls]<sub>2</sub> to [Junnus]<sub>3</sub>]<sub>4</sub>**, when **[[I]<sub>1</sub>** still agree to explain [this matter]<sub>2</sub> to [you]<sub>3</sub>]<sub>4</sub>.'

The message in which this token appears was written as a reply to a discussion participant whose first name is *Junnu*. Thus, *sä* 'you' in *kun suostun sulle tätä asiaa edelleenkin selvittämään* 'when I still agree to explain this matter to you' and *Junnu* in *syyttää helmiä Junnuille* 'throw pearls to Junnus' have the same indexical feature, marked by the subscript number 3 in the text and in Figure 134 below.

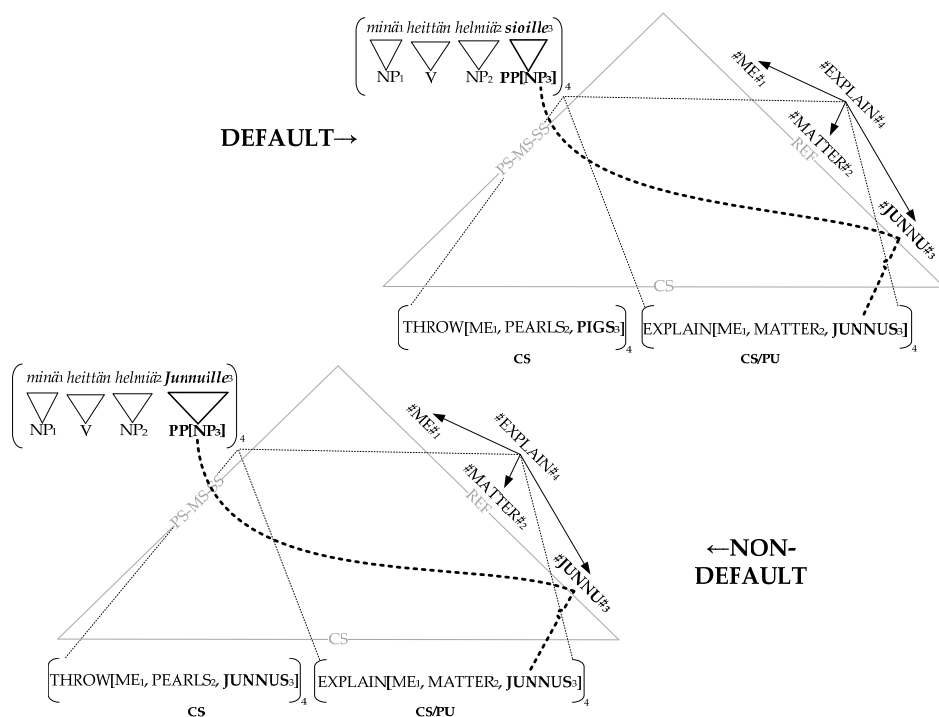


Figure 134 Analysis of non-default linking to phonology in (293)

In (294) below the NP<sub>SUBJ</sub> constituent of the intransitive construction NP<sub>SUBJ</sub>[N<sub>1</sub>] V<sub>INTR</sub> PP<sub>ADJUNCT</sub>[NP[N<sub>2</sub>]] corresponds to the phonological form *kaunorunollinen ironiani* 'poetic irony<sub>1SG</sub>; my poetic irony' instead of the default *helmi* 'pearl':

(294) *Ehdin jo pelätä, että [[kaunorunollinen ironiani]<sub>1</sub> oli mennyt täysin [sioille]<sub>2</sub>]<sub>3</sub>.*<sup>415</sup>

<sup>415</sup> GG: sfnet.harrastus.veneet/msg/e7657ef082d68740, Feb 10 2005

lit. 'I already feared that [[my poetic irony]<sub>1</sub> had gone completely to [pigs]<sub>2</sub>]<sub>3</sub>.'

Non-default correspondence in the above example can be explained by the fact that *kaunorunollinen ironiani* 'poetic irony<sub>1SG</sub>; my poetic irony' refers to the #CONTENT# in the #COMMUNICATIVE EVENT# (see Figure 135 below).

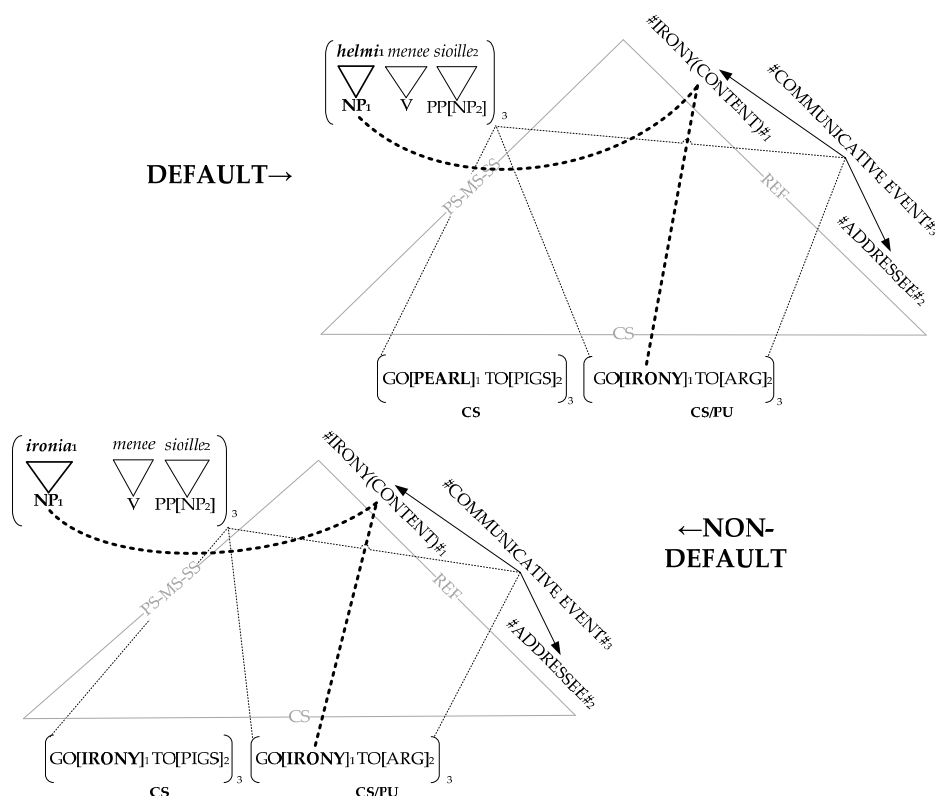


Figure 135 Analysis of non-default linking to phonology in (294)

#### 5.4.2.2 Referentially licensed non-default verbs

A PU can be used to conceptualize different situations, e.g. communication in (295) and (296); caused change of location in (297) and (298) or caused transfer of possession in (299). Conceptual structures of non-default verbs presented here are identical with the argument structure of the main conceptual clause in the CS/PU corresponding to the intended #SITUATION#/PU.

- (295) *Haukkumalla kirjoittajaa siaksi, jolle et kuitenkaan helmiäsi suostu paljastamaan.*<sup>416</sup>  
lit. 'By calling the author a pig, to which you will not **reveal** your pearls.'
- (296) *Kun vastassa on ymmärtämättömyys, kokemuksen puute ja haluttomuus tarkastella asioita avoimesti ja ennakkoluulottomasti, perustelun 'helmiä' ei näille saparohännille kannata kädestään näyttää.*<sup>417</sup>  
lit. 'When faced with lack of understanding, lack of experience and unwillingness to look at things openly and open-mindedly, it's pointless to **show** 'pearls' of explanation to these short tails.'
- (297) *Jenkit keräsivät suomen koululaisille paketteja joita jaettiin kansakuoluissa. Minäkin sain paketin se sisälsi lasihelmiä ja pulverihammastahnaa rasian. [...] Niistä lasihelmistä vuosien mittaan on minulle selvinnyt että jenkit lähettivät lasihelmiä sioille.*<sup>418</sup>  
lit. 'The Yanks gathered packages for Finnish schoolchildren that were distributed in elementary schools. I got a package too it contained glass beads and a box of powder dentifrice. [...] Over the years I learned about these glass beads that the Yanks **sent** glass beads to pigs.'
- (298) A: Voit lähettää mulle loppumattomasti niitä 12 euron osajakirvesmiehiäsi.  
B: Sulla on sen verran "ahdistusta" käyttäytymisessä, että ne ei sulle tai sun välittämänä kauaa työskentelis, joten ei suotta **laiteta** "helmiä sioille".<sup>419</sup>  
lit. 'A: You can send me endlessly these 12-euro master carpenters of yours.  
B: You have so much "anxiety" in your behaviour that they would not work long for you or through you, so let's not **send** "pearls to pigs" in vain.'

Let us look more closely at the non-default verb *myydä* 'sell' in (299) below:

- (299) A: Ja koska [**Porsche**]<sub>1</sub> on jättänyt itselleen oikeuden valita [**kyseisen mallin**]<sub>2</sub> [**ostajat**]<sub>3</sub>, [**tilaajan**]<sub>3</sub> on pakko olla [**pidemmän linjan possumies**]<sub>5</sub> (vaiko peräti -nainen näin tasa-arvon aikoina?).<sup>420</sup>  
B: Tottakai. Miksi [**myydä** helmiä<sub>2</sub> sioille<sub>3</sub>]<sub>4</sub>...vai pitäisikö sanoa possumilles<sub>3</sub>.<sup>421</sup>  
lit. 'A: And, since [**Porsche**]<sub>1</sub> has reserved the right to choose [**buyers**]<sub>3</sub> for [the model in question]<sub>2</sub>, the [**orderer**]<sub>3</sub> has to be [an old-time **piggy-man**]<sub>5</sub> (or even a woman, in these times of equality?)

<sup>416</sup> GG: sfnet.urheilu/msg/d49fa37eeb4a3547, Dec 11 2000

<sup>417</sup> GG: sfnet.keskustelu.psykologia/msg/2b2aa217882b69bc, Mar 28 1999

<sup>418</sup> GG: sfnet.keskustelu.maanpuolustus/msg/e9b098106350373b, Feb 15 2005

<sup>419</sup> GG: sfnet.keskustelu.rakentaminen/msg/716d8493cda221da, Sep 7 2006

<sup>420</sup> GG: sfnet.harrastus.autot/msg/1bf3c08e2f82b6a8, Feb 28 2002

<sup>421</sup> GG: sfnet.harrastus.autot/msg/599b744db8a1390c, Feb 28 2002

B: Sure. Why [sell pearls<sub>2</sub> to pigs<sub>3</sub>]<sub>4</sub> ... or should I say **to piggies**.'

The intended referent of the PU in this example is the situation (marked with a subscript index 4) of Porsche (marked with a subscript index 1) selling specific car models (index 2) only to their long-time customers (index 3). Thus, non-regular customers would not be qualified as adequate buyers of this car model and therefore are the referent of *siat* 'pigs'. *Possu* 'piggy, piglet' is in fact a Finnish slang word for Porsche sports cars, which explains the use of the epithet *possumies* 'lit. piggy-man; Porsche-man' denoting a Porsche-owner. The semantic proximity of *possu* 'piggy, piglet' and *sika* 'pig' serves as a basis for the pun *vai pitäisikö sanoa possuille* 'or should I say **to piggies**'. The analysis of the example (299) is presented in Figure 136 below.

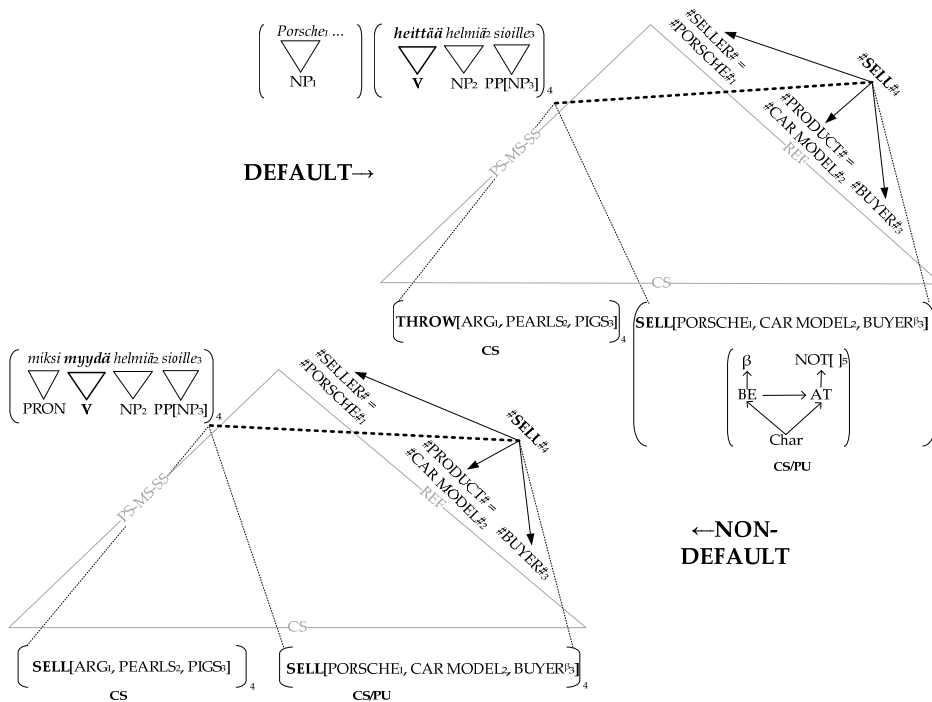


Figure 136 Analysis of non-default linking to phonology in (299)

### 5.4.3 Phraseological constituent isolation (PCI)

Phraseological cohesion, i.e. an association that exists between lexical items created by their co-occurrence within the same PU, as well as an association between the items and the PU itself, discussed in Section 5.2.1 of the present chapter, can result in the ability of default lexical items –

noun constituents of a PU – to occur separately, without any immediate structural relation to other constituents, as a part of some larger structures that do not belong to the group of (more or less) conventionalized constructions of the PU's construction family. In Petrova (2007b) I have labelled this phenomenon as phraseological constituent isolation (PCI) and such items as isolated phraseological constituents (IPC), e.g. in (300) below both *NP<sub>PL</sub>{helmi 'pearl'}* and *NP<sub>PL</sub>{sika 'pig'}* occur twice in isolation from each other:

(300) A: *Jeesus Nasaretilainen taisi sanoa saman asian jotenkin, että --- hmmm -- ainakin hän sanoi, ettemme saisi heittää sioille helmiämme.*<sup>422</sup>

B: *Muttä täällä newsseissä sitä vain roikutaan, heh? Eikä [sioissakaan]<sub>IPC</sub> mitään vikaa ole. (Hengellisiä) [helmiä]<sub>IPC</sub> kun on loputtomasti tarjolla. Toisaalta [siat]<sub>IPC</sub> haluaisivat ennemmin ruokaa, koska siitä on heille enemmän hyötyä. Mutta [helmet]<sub>IPC</sub> voivat ne tappaa joutuessaan henkireikään.*<sup>423</sup>

lit. A: 'I think that Jesus from Nazareth said the same thing somehow, that --- hmmm – at least he said that we are not allowed **to** throw our pearls to pigs.'

B: 'But people are just hanging out here in newsgroups, heh? There's nothing wrong with [**pigs**]<sub>IPC</sub> either. For (spiritual) [**pearls**]<sub>IPC</sub> are in endless supply. On the other hand, [**pigs**]<sub>IPC</sub> would rather like food, because it is of more benefit to them. But [**pearls**]<sub>IPC</sub> can kill them if they get into the air hole.'

As far as I know, PCI has never been examined as a special case of idiom variation. For instance, Naciscione (2001: 108) treats it as a part of a larger phenomenon, which she calls phraseological allusion and defines as "an implicit mental reference to the image of a phraseological unit which is represented in discourse by one or more explicit image-bearing components". Thus her main criteria for allusion is that the PU does not appear in its full base form, while the structural relations of the remaining constituents are not taken into account. As far as my definition is based primarily on mutual structural relation of constituents, I prefer to make a distinction between cases of a true syntactic isolation from other PU constituents and cases where default items occur within the syntactic structure of the same clause and, thus, are both structurally and cohesively related to each other at the sentential level, although the

<sup>422</sup> GG: sfnet.keskustelu.uskonto/msg/89bd0bbb59e2f78e, Dec 22 1997

<sup>423</sup> GG: sfnet.keskustelu.uskonto/msg/4d2c63d56482e68f, Dec 24 1997

structure itself may not be a conventionalized construction of the family in question, e.g. (301) below:

- (301) *Helmet eivät kelpaa ravinnoksi sioille, ei vaikka kuinka kimaltelisivat. Ja päinvastoin.*<sup>424</sup>  
 lit. 'Pearls are unfit for pigs' nutrition, no matter how they shine. And vice versa.'

Isolated constituents can occur in the context preceding the phraseological unit, subsequent to it, or without immediate reference to the PU's local linguistic context. On the basis of this distinction three different types of PCI can be described. The first type involves presence of the PU in the same context with an IPC that refers to it either anaphorically, as in (302) below, or cataphorically, as in (303) below. Their relation to each other is cohesive, not structural:

- (302) *Toisaalta minun uskovana on varottava esim. rikkomasta Jeesuksen käskyä [olla antamatta helmiä sioille]<sub>PU</sub>. Täälläkin liikkuu varmasti noita [sikoja]<sub>IPC</sub>...*<sup>425</sup>  
 lit. 'On the other hand, as a believer I have to beware of e.g. disregarding the command of Jesus [not to give pearls to pigs]<sub>PU</sub>. Those [pigs]<sub>IPC</sub> are certainly moving around here as well...'
- (303) *Sivistys, poika hyvä, on vähän niin kuin [helmet]<sub>IPC</sub>: [sitä ei sioille heitellä]<sub>PU</sub>.*<sup>426</sup>  
 lit. 'Culture, my dear boy, is a bit like [pearls]<sub>IPC</sub>: [one does not throw it to pigs]<sub>PU</sub>.'

In the second type several IPCs occur in the same context with no reference to the PU. Cohesion between the IPCs and the missing PU is established by virtue of the phraseological cohesion and is supported by the discourse cohesion, e.g. (304) below:

- (304) *[Sikoja]<sub>IPC</sub> kyllä telkusta löytyy, [helmiä]<sub>IPC</sub> saa hakea.*<sup>427</sup>  
 lit. 'Indeed [pigs]<sub>IPC</sub> can be found on the telly, [pearls]<sub>IPC</sub> have to be searched for.'

The third type includes occurrences of a single IPC, while neither the PU nor any of the other default lexical constituents are present in the same context. Again, connection between the IPC and the PU is established by virtue of phraseological cohesion; furthermore, a particularly important role is played by the discourse cohesion:

<sup>424</sup> GG: sfnet.keskustelu.kielipolitiikka/msg/c855b46a87a90173, Mar 5 2003

<sup>425</sup> GG: sfnet.keskustelu.evolutio/msg/67db272e6d312330, Mar 5 2003

<sup>426</sup> GG: sfnet.keskustelu.talous/msg/f127e16009910ef3, Mar 11 2005

<sup>427</sup> GG: sfnet.viestinta.tv/msg/84a0620b35a646b4, 16 Feb 2004

- (305) *Mutta ihan oikeasti; [jos [sinä<sub>1</sub> olet<sub>2</sub> uskontovastainen<sub>3</sub> eikä<sub>4</sub> sinua<sub>1</sub> [aidosti kiinnostaa]<sub>5</sub> kuulla<sub>6</sub> [mitä buddhalaisuus pitää sisällään]<sub>7</sub>]<sub>8</sub>, niin [[voi[mme]<sub>1,9</sub>]<sub>10</sub> lopettaa]<sub>11</sub> [tämän keskustelun]<sub>12</sub> sitten tähän]<sub>13</sub>]<sub>14</sub>. Tämä<sub>14</sub> on tradition määräämä juttu; ei minun keksintöni. [Jeesuksella<sub>15</sub> oli<sub>16</sub> muuten tähän]<sub>14</sub> [oiva sanonta]<sub>17</sub>]<sub>18</sub> – [siinä<sub>17</sub> oli<sub>19</sub> [joitain [helmiä]<sub>IPC</sub> ja sen semmoisia.]]<sub>20</sub>]<sub>21</sub><sup>428</sup>*

lit. 'But really; [if [you<sub>1</sub> are<sub>2</sub> antireligious<sub>3</sub> nor<sub>4</sub> are you<sub>1</sub> [genuinely interested]<sub>5</sub> to hear<sub>6</sub> [what Buddhism contains]<sub>7</sub>]<sub>8</sub>, then [we<sub>1,9</sub> can<sub>10</sub> end<sub>11</sub> [this conversation]<sub>12</sub> here]<sub>13</sub>]<sub>14</sub>. This<sub>14</sub> is a thing set by a tradition; not my invention. By the way, [Jesus<sub>15</sub> had<sub>16</sub> [an excellent saying]<sub>17</sub> about this]<sub>14</sub>]<sub>18</sub> – [there<sub>17</sub> were<sub>19</sub> [some [pearls]<sub>IPC</sub> and stuff like that]<sub>20</sub>]<sub>21</sub>.

In the above example, the IPC *NP<sub>PTV</sub> PL{helmi 'pearl'}* occurs with no immediately available reference to any other default constituent or any construction of the *HHS* construction family within the same context. Since only one default lexical item is present in the text, the process of PU retrieval and interpretation by the reader can be considerably impeded. The reader is given a possibility to make an inference that the isolated constituent does in fact refer to some PU due to the metarepresentation (pure quotation or mention; Noh 2000) of the PU by means of an appellative *sanonta* 'saying'. In case the reader is familiar with the PU's Biblical origin (see Section 5.2.3 for the discussion on textual dependence in PUs), (s)he can also establish a connection between this knowledge and the evidential *Jeesuksella* 'Jesus<sub>SADE</sub> SC' in the passage [*Jeesuksella*]<sub>source</sub> *oli muuten tähän oiva [sanonta]*<sub>metarepresentation</sub> '[Jesus] had an excellent [saying] about this'. However, the same evidential could in principle refer to another saying by Jesus, which also contains the word *helmiä* 'pearls', namely the "Parable of the Pearl of Great Price" presented in (306) below:

- (306) *Taivasten valtakunta on myös tällainen. Kauppias etsi kauniita helmiä. Kun hän löysi yhden kallisarvoisen helmen, hän myi kaiken minkä omisti ja osti sen.* (Matt. 13: 45-46)

'Again, the kingdom of heaven is like a merchant seeking beautiful pearls, who, when he had found one pearl of great value, went and sold all that he had and bought it.'

However, the above quotation did not originate any PU in the Finnish language. There is also a mismatch between the meaning of the parable in (306) and the statement made by the writer in (305) above about the existing referential connection between the *sanonta* 'saying', on the one hand, and the content of logical implication marked with subscript index

<sup>428</sup> GG: sfnet.keskustelu.uskonto/msg/2c7403e0ea807863, May 1 2002

14, on the other. The intended PU's conceptual structure has, on the contrary, several important connections to this content. The formal analysis of (305) is presented in Figure 137 and in Figure 138 below there is a formal analysis of the intended PU, where parts are co-indexed with elements of the text in (305). Below is the list of the most relevant parts of the context numbered according to their index in (305) and in Figure 137:

1 = YOU – deictic referent of the ADDRESSEE

3 = ANTIRELIGIOUS – PROPERTY of the ADDRESSEE

7 = CONTENT of the COMMUNICATIVE EVENT

8 = A complex proposition, conceptualization of ADDRESSEE's INADEQUACY as a recipient of the CONTENT (i.e. being an antireligious person, who is not genuinely interested in hearing what Buddhism contains)

1, 9 = WE (in Finnish expressed by the first person plural ending of *voimme* 'can<sub>PL1</sub>; we can') deictic reference to both participants in the event – the SPEAKER and the ADDRESSEE

12 = COMMUNICATIVE EVENT

13 = Proposition expressing the possibility for the COMMUNICATIVE EVENT to be ended by its participants

14 = the logical implication "if ... then ..." postulated between propositions 8 and 13

15 = JESUS – "owner" of the "author's rights" to the PU

17 = SAYING – metarepresentation of the PU co-referential with 14

18 = Evidential establishing the authorship of PU

20 = PEARLS – the isolated constituent of the PU (IPC)

21 = LOCATION of the IPC in the PU

Formal analysis of the intended PU (represented here by the transitive verbal construction) in Figure 138 shows that a proposition indexed in Figure 137 with 8, which characterizes the ADDRESSEE as an inadequate recipient of the further argumentation on the matter of Buddhism, in the CS/PU corresponds to the property INADEQUATE assigned to the thematic argument co-referential with the *NP<sub>PL</sub>{sika 'pig'}*. As has been previously mentioned in Section 5.4.1 above, the default semantic fields in the main conceptual clause of the CS/PU are Comm (Communicative) and Cogn-Perc (Cognitive-Perceptual), while the proposition itself by default encodes a communicative event. In the CS of the discourse fragment (305) presented as several propositions in Figure 137, this communicative event corresponds to the thematic argument [CONVERSATION]<sub>12</sub>.



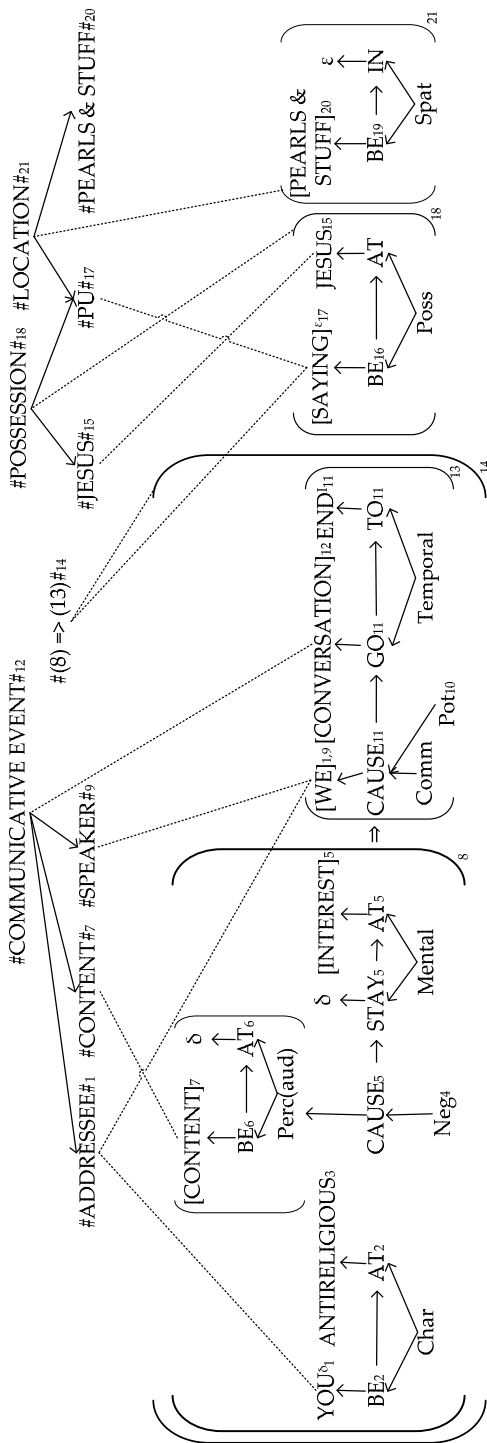


Figure 137 Analysis of discourse-cohesive elements in (305) above



However, there are some problematic aspects of the utterance meaning that cannot yet be successfully resolved in the formal description of its CS. The first one is the formal representation of logical relations between complex utterances. In Figure 137 above the technical solution was to connect two parts of the implication by a “=>” sign. The problem itself is rather complex and lies beyond the scope of the present study. The other problem concerns pragmatic implicatures of the modal auxiliary *voida* ‘can’ as in *[[voi[mme]<sub>8</sub>]<sub>10</sub> lopettaa<sub>11</sub> [tämän keskustelun]<sub>12</sub> sitten tähän]<sub>13</sub>]<sub>14</sub>* ‘[we<sub>8</sub> can<sub>10</sub> end<sub>11</sub> [this conversation]<sub>12</sub> here]<sub>13</sub>]<sub>14</sub>’. The default interpretation of this Finnish modal verb is a dynamic possibility, although in certain pragmatic uses (where this auxiliary often appears in the conditional mood) it can also express directive modality with such functions as suggestion (*Voisit yhdistää nämä kaksi asiaa.* ‘You could connect these two things.’), offer (*Voisinko auttaa?* ‘Could I help?’), permission (*Voitte ottaa lapsetkin mukaan.* ‘You can bring the children too.’), and request (*Voisitteko olla hiljempää?* ‘Could you be quieter?’) (Hakulinen et al. 2004). Thus, the unmarked paraphrased meaning of the above-mentioned sentence would be ‘external circumstances allow us to end this conversation here’. However, considering the fact that the writer refers to the PU, where the communicative event is selected as a scope for modality of inexpediency (see Chapter 3 Section 3.4.3.6.2 and Section 5.3.2 of the present chapter), use of the verb *voida* ‘can’ in this particular context could have other implications than its unmarked meaning: it could implicitly express the writer’s suggestion that the conversation should be ended. Nevertheless, subjective modality and pragmatic polyfunctionality of modal verbs (Laitinen 1993, Coates 1990, van der Auwera 1999) teeter on the edge of the semantic/pragmatic interface, which is a rather complex issue and deserves a thorough investigation (e.g. Larjavaara 2007, Turner 1999). Such uses of modal verbs may pertain to different conventional stylistic, rhetoric, politeness or face-saving strategies, but it is a separate discussion topic, which goes outside the scope of this thesis. In order to be faithful to one’s guidelines of formal semantic description one needs to keep it apart from subjective interpretations.

In order to successfully interpret the fragment, the addressee has to access the PU, which of course presupposes that the given unit has been previously learned and is present in the subject’s long-term memory. Jackendoff (2002: 210) mentions four components of lexical access:

activation, binding, integration and resolution. Jackendoff's (2002b: 153) notion of a lexical item denotes any item stored in the lexicon, i.e. in long-term memory, and includes, besides words, items smaller than words (morphology) and larger than words (idioms). If the input from reading interface would have contained a full form of this PU, its activation in long-term memory could have been accomplished by a call from phonological working memory to the unit's default phonological structure. However, in the case of the example in (305) above the input to phonological working memory contains only a single IPC *helmiä* 'pearl<sub>PL TV PL</sub>'. As a result, the PU itself has to be activated via the internal cohesive links that exist between its lexical constituents and the external discourse-cohesive links that has been described here.

## 5.5 Summary

In this chapter I have presented the Tiernet model of PU variation and formulated its main principles. The main advantage of this model is that it allows one to abandon the hierarchical and categorical schemes of variation and concentrate specifically on linking between different levels of representation. I have argued that instead of being presented as nodes in a hierarchical network, which inevitably turns them into large chunks of information, constructions can themselves be treated as networks, i.e. a totality of linking patterns, licensed by them. Thus, a PU's construction family is a vast network formed by the default pattern as well as each and every non-default realization. When several patterns are incorporated into a single network, each construction represents a unique combination of nodes and links, and, what is particularly important, none of them is treated as a derivation from some base form, or a more abstract construction, as has been done in previous models of variation. Similarities between constructions in a family can be described as overlapping nodes of a particular tier. Thus, both similarities and differences are always tier-specific.

Construction family variation of the Finnish PU *X HEITTÄÄ HELMIÄ SIOILLE* | *NP<sub>SUBJ</sub>[N{X}] V{heittää 'throw'} NP<sub>OBJ</sub>[N<sub>PL</sub>{helmi 'pearl'}] PP<sub>ALL</sub>[N<sub>PL</sub>{sika 'pig'}]* | lit. 'X throws pearls to pigs', id. 'X causes a transfer of some entity, evaluated by the speaker as good, to a recipient, evaluated as inadequate in some way', also referred to as the *HELMIA SIOILLE* construction family (or *HHS*), has been analyzed in this chapter from two

different perspectives – the structure of the PU and the discourse context. I have developed the notion of phraseological cohesion (Naciscione 2001) into a more elaborate, explicit and formally plausible association between the parts of a PU's network structure, on the one hand, and between the latter and the discourse context, on the other. It is true that in many cases both the structure and the discourse act as underlying sources of variation, but the model is able to handle any of such cases. An important point that has to be made here is that an explicit detailed description of PU structure opens a whole new world of possibilities for the study of PU variation and even provides fundamentally new tools for textual analysis.

## 6. Conclusion

In Section 1.2 of the Introduction I mentioned the development of a theoretical and descriptive model of phraseological units' (PU) structure and variation as the primary objective for this study. In order to build such a model and to show how it works as a system it was necessary to undertake an in-depth usage-based analysis of one single PU, which is quite an unorthodox approach compared to the prevailing trends in mainstream research on idioms and phraseology. Another important difference is the initial strive for an explicit formal description, which is very seldom a central issue in research on idioms. A critical overview of some of the previous approaches to the analysis of PU structure and variation was presented in Chapter 2 and Chapter 4 of this book. Chapters 3 and 5 present the main bulk of the research carried out. It has been argued that phraseological units (PUs) are complex structures that should be described simultaneously at several different levels (phonology, morphology, syntax, CS, reference). Thus, it was necessary to develop a model that could integrate formal description at all these levels. Conceptual-semantic theory and especially the Tiernet model (Nikanne 2002, 2006) proved to be a very effective framework for this task.

Although the notion of construction family, which goes back to the Wittgensteinian idea of family resemblance, has been adopted from construction grammars (Goldberg 1995, 2006; Croft 2001), the construction family network model of PU variation presented in this book differs from construction family networks in construction grammars in a very important respect – in the present model variation is not treated as being based on derivation and inheritance from more abstract levels of representation. Instead, it is described in terms of linking constraints, which, together with the connectionist architecture of the model, is in accord with Jackendoff's view of linguistic competence. I have argued that PU variability or the lack of such can be examined on a node-to-node basis and that the system of linking values (such as fixed, default or non-default) can be used to formalize constraints. Since the present model of PU variation is initially usage-based, linking values at each level are to be determined empirically. Different text types may produce somewhat different quantitative results, which can only affect linking values, but the core architecture (the essence) of the PU's structure remains intact. Thus, the efficiency of this model lies in its ability to describe idiom variation at

all relevant levels without losing connection to the larger linguistic structure. On the other hand, there is no more need for categorization of variation, since I am able to zoom on a specific link (or combination of links) at a time and examine variation in that segment of network.

In this book I have also addressed the problem of the base form of PUs – a highly relevant notion for the study of their variation and variability. The problem of the base form approach is that there are neither clear criteria of how such a form could be determined, nor adequate solutions to the methods of its representation that could account for possible variation. If all three major criteria (syntactic construction, constituent morphology and constituent phonology) are to be simultaneously met in order for a PU token to be considered a base form token in a strict sense it is not guaranteed that the resulting form will be the most frequent form in a corpus, as it actually happens in the case of *HELMIÄ SIOILLE* constructions. In this respect, the default value determined for each particular link separately seems to be a much more flexible solution. Phonological (lexical) and morphological defaults for constituents in the *HELMIÄ SIOILLE* construction family have been calculated. The possibility of calculating purely syntax defaults is not ruled out; however, it is a much more complicated task, since many different links are involved in a complex syntactic structure. One still needs more precise criteria of default linking value compared to free, unrestricted mapping, strong vs. weak default patterns, etc. Such criteria have to be based on robust statistical methods.

Several theoretical and practical problems had to be tackled in order to arrive at a plausible model of the PU's structure. An important distinction between the current approach to the PU's semantic structure and the majority of previous approaches to idiomaticity is my rejection of the distinction between literal meaning and idiomatic meaning as two essentially different types of meaning. I have argued that there is nothing "idiomatic" about the idiom's actual meaning per se, since both the "literal" and the "idiomatic" semantic structures are formed using the same primitive units according to the same conceptual formation principles. Thus, it is not in the conceptual structure that one has to look for idiosyncrasies, but in its correspondence to another representational module (syntax). Idiosyncratic connection between semantic and syntactic structures can only be postulated if one accepts that there are linguistic phenomena where linking between these two modules is

governed by some regular principles. Here I agree with Nikanne (2005a, 2008a), who emphasizes the importance of regular structures and at the same time does not downplay irregular ones. Thus, instead of the term literal meaning I have used the notion of the CS (conceptual structure) as a structure, which is mapped to syntactic module according to regular principles of syntactico-semantic linking. Instead of the term idiomatic meaning I have introduced the notion of CS/PU - a structure, whose idiosyncratic linking to syntax is licensed by a PU. It has been argued that it is both psycholinguistically and theoretically plausible to assume the presence of both structures in semantically ambiguous PUs.

Another important problem is the notion of analyzability. It has been previously described as a kind of isomorphism – a one-to-one correspondence between the formal and the semantic structures of a PU. I have argued that there is no possibility to link CS/PU thematic arguments directly to syntax, because the syntactic roles are already assigned by the rule-based CS. Instead, there is a more plausible way to treat analyzability as referentiality of constituents licensed by the PU (#SITUATION#/PU referentiality), by assuming that linking of the CS/PU to the level of syntactic representation is mediated by the assignment of #SITUATION#/PU referential indices to the syntactic constituents of the PU. Although the matter requires a more extensive study of a greater number of PUs than could be done within the scope of this study, there is good reason to believe that analogy plays an important role in this process. I.e., given the identity of f-chains in both conceptual structures, one can expect a high probability that syntactic constituents corresponding to thematic arguments in the CS will be mapped onto the same #ENTITY#/PU indices in the #SITUATION#/PU as arguments selected by the same thematic functions in CS/PU.

Yet another advantage of the conceptual-semantic model is that with the help of its formalism it is possible to describe semantic structure of PUs by providing one sufficiently underspecified semantic representation which is compatible with all possible contextual meanings one might encounter in empirical data. As has been demonstrated with the usage-based analysis of the Finnish PU *X HEITTÄÄ HELMIÄ SIOILLE* | *NP<sub>SUBJ</sub>[N{X}] V{heittää ‘throw’} NP<sub>OBJ</sub>[N<sub>PL</sub>{helmi ‘pearl’}] PP<sub>ALL</sub>[N<sub>PL</sub>{sika ‘pig’}]* | lit. ‘X throws pearls to pigs’, id. ‘X causes a transfer of some entity, evaluated by the speaker as good, to a recipient, evaluated as inadequate in some way’,



the contextual-semantic variation in PUs is to a large extent possible due to the ability of the semantic field tier (S-tier) – an abstract organizing system for concepts – to have different realizations. Semantic fields certainly need much more thorough study than was possible to accomplish within the limits of the present dissertation.

The other essential underlying source of variation is reference and textual cohesion. Until present, Conceptual Semantics has been mostly a theory concerned with description of linguistic structure, while its relations to the context have remained largely unexplored. In this respect the present study is truly innovative. The context of computer-mediated communication proved to be a fertile ground for the study of the PU's text-binding function. It has been demonstrated how the semantic disambiguation and specification of the PU can be achieved by resolving referential indeterminacies and how the relevant textual elements can be determined and formalized, i.e. explicitly marked in the text. As long as I am concerned with linguistic structure alone, the theoretical and methodological basis of Conceptual Semantics is quite sufficient for this kind of analysis, but as soon as one starts shifting the focus from structure towards communicative aspects of PU use, one notices that there is still much to be done. The explicit formal analysis of contextual, discourse-pragmatic and social aspects of PUs within the conceptual-semantic framework definitely deserves further development. Since the Tiernet model is initially open, nothing prevents one from building on it. Jackendoff (2007) and Paulsen (2011) have already discussed possibilities to include fragments of social knowledge into linguistic description. There could also be some good use for an interface which encodes attitudes and dialogicity. Since Conceptual Semantics yet lacks its own tools for discourse-pragmatic analysis, it is definitely worthwhile exploring in the future whether Frame Semantics or any other theory has something to offer in order to accommodate such features of PUs into the Tiernet model. However, mutual compatibility of theories, including basic assumptions about language that they make and methodological guidelines that they follow, has to be taken into consideration before any attempts to combine their notions and metalanguage can be made.

## Svensk sammanfattning

Fraseologiska enheter är konventionella och relativt stabila flerordiga uttryck som uppvisar olika slags strukturella oregelbundenheter. De utgör en väsentlig del av vår kunskap om språket och erbjuder rikligt med möjligheter att förbättra vår förståelse om det sätt på vilket komplexa, abstrakta strukturer och principer som bygger upp människornas språkförmåga fungerar. Att studera dessa uttryck är både givande och utmanande. Trots att de är frasenheter, är de – precis som ord – lagrade i språkbrukarnas minnen. Trots att de flesta fraseologiska enheterna följer allmänna syntaktiska regler, är länkarna mellan syntax och semantisk struktur som de skapar oregelbundna. Deras förmåga att genomgå variation kan vara idiosynkratiskt begränsade till så hög grad att de är fullständig orubbliga. Trots detta tillåter många av dem variation, t.o.m. i sådan utsträckning att gränserna mellan en variant och en standard form suddas ut, och istället för ett uttryck råkar vi plötsligt på en familj av flera besläktade konstruktioner. De fraseologiska enheternas lättigenkännlighet och deras samtidiga oerhört stora variation väcker frågan om hur dessa uttryck egentligen ska vara representerade i språket. Syftet med denna avhandling är: 1) att utföra en kritisk granskning av olika tidigare tillvägagångssätt att beskriva fraseologiska enheter i syfte att klargöra vilka svårigheter deras struktur, variation och variabilitet erbjuder för den lingvistiska teorin samt 2) att presentera ett alternativt sätt att beskriva dessa variabla uttryck.

Fraseologiska enheter är komplexa strukturer som ska samtidigt beskrivas på olika nivåer (fonologi, morfologi, syntax, konceptuell struktur, referens). Således var det nödvändigt att utveckla en systematisk modell som kunde integrera en formell beskrivning på alla dessa nivåer. Den teoretiska och metodologiska referensramen för modellen som beskriver fraseologiska enheter i denna avhandling utgörs av Konceptuell Semantik (Jackendoff, Nikanne, Pörn, Paulsen) och i synnerhet av Tiernetmodellen (Nikanne). Den sistnämnda visade sig vara ett mycket effektivt deskriptivt redskap. Den grundläggande idén bakom Tiernet utgörs av ett formellt analytisk sätt att beskriva språket med. Man använder enkla moduler som är sammankopplade av förbindelselänkar. I avhandlingen analyseras de fraseologiska enheternas struktur i form av ett nätverk med tillämpning av Tiernetmodellen.

Wittgensteins idé om familjelikheter ligger i bakgrunden för analysen av variation där begreppet konstruktionsfamilj tas i bruk. I konstruktionsfamiljnätverksmodellen, som beskriver idiomvariation och som presenteras i denna avhandling, behandlas inte variationen med avledningar från en basform som utgångspunkt. Till skillnad från tidigare fraseologiska och konstruktionsgrammatiska modeller ses variationen av idiom inte heller som ett arv från mer abstrakta nivåer eller som en kategorisering av olika

variationstyper. Istället beskrivs variationen som ett system av länkvärden som används för att formalisera restriktioner bland konstruktionerna. Dessa värden beräknas för varje nivå på basis av en kvantitativ analys av ett korpusmaterial av 588 förekomstexempel av ett idiom som återfinns i diskussionsgrupper i det finska Usenet. Modellen bygger på en fördjupande, språkbrukbaserad analys av en fraseologisk enhet – det finska idiomet *X HEITTÄÄ HELMIÄ SIOILLE* id. 'X kastar pärlor åt svin', id. 'X förorsakar överföring av någon enhet, evaluerad av talaren som bra, till en mottagare, evaluerad av talaren som inadekvat'. Detta ska inte låta som något begränsande – under loppet av denna studie hänvisas också till många andra fraseologiska enheter, även om inte lika ingående som till det ovannämnda idiomet. Analysen av denna enhet, för vilken ett omfattande korpusmaterial insamlats, leder till att många andra fenomen oundvikligen dras in.

Flere teoretiska och metodologiska problem måste lösas innan en plausibel modell av den fraseologiska enhetens semantiska struktur kunde konstrueras. En väsentlig skillnad mellan den syn på denna struktur som presenteras i detta arbete och de flesta tidigare sätt att behandla idiomatiskhet är att ingen gräns dragits i detta arbete mellan bokstavlig betydelse och idiomatisk betydelse som om dessa vore två väsensskilda typer av betydelse. Argumentet är att idiomets betydelse inte är idiomatisk per se, eftersom både den regelbundna och den idiom-licenserade konceptuella strukturerna är uppbyggda av samma primitiva enheter och enligt samma principer. Det är alltså inte i den själva konceptuella strukturen vi ska leta efter idiosynkratiska drag, utan i korrespondenslänkar mellan den och syntaxen. Ett annat viktigt problem är begreppet analyserbarhet av idiomets semantiska struktur. Analyserbarheten har tidigare beskrivits som ett slags isomorfi eller en-entydig korrespondens mellan den formella och den semantiska strukturen i ett idiom. I denna avhandling framgår det dock att det är omöjligt att direkt länka tematiska argument i den idiom-licenserade konceptuella strukturen till syntaxen, eftersom syntaktiska roller redan tilldelats idiomets av den regelbundna konceptuella strukturen. I avhandlingen presenteras ett alternativt sätt att behandla analyserbarhet: Den indirekta korrespondensen mellan den idiom-licenserade konceptuella strukturen och syntaxen förmedlas av den referentiella strukturen. Referentiell bindning betraktas i denna studie som en viktig variationsmekanism. Kontexten av datorförmedlad kommunikation visade sig vara en fruktbar grund för undersökning av idiomets textbindande funktion. I denna avhandling demonstreras hur den semantiska disambigueringen och specificeringen i en fraseologisk enhet kan uppnås genom upplösning av den referentiella obestämdheten och hur de relevanta textuella elementen kan formaliseras, d.v.s. tydligt markeras i texten.

## References

- Abeillé, A. 1995, "The flexibility of French idioms : a representation with Lexicalized Tree Adjoining Grammar" in *Idioms : structural and psychological perspectives*, eds. M. Everaert, E. van der Linden, A. Schenk & R. Schreuder, Lawrence Erlbaum Associates, Hillsdale, New Jersey, pp. 15-42.
- Ahmanova, O.S. 1966, *Slovar' lingvističeskikh terminov*, Sovetskaâ ènciklopediâ, Moscow.
- Aksamitov, A.S. 1965, "Struktura ideologičeskogo slovarâ belorusskoj frazeologii" in *Voprosy frazeologii : frazeologija masalalari*, ed. L.I. Rojzenzon, Nauka, Taskent, pp. 10-14.
- Albright, A.C. 2002, *The identification of bases in morphological paradigms*, PhD thesis, University of California, Los Angeles, CA.
- Alefirenko, N.F. & Valjuh, Z.O. 1990, "Teoretičeskie osnovy frazeologičeskoj reprezentacii formal'no-soderžatel'noj variantnosti frazeologizmov", *Problemy russkoj i obšej frazeografii : mežvuzovskij sbornik naučnyh trudov*, Novgorod, pp. 27-28.
- Amosova, N.N. 1963, *Osnovy anglijskoj frazeologii*, Izdatel'stvo Leningradskogo universiteta, Leningrad.
- Apresjan, J.D. 1995, *Izbrannye trudy. Integral'noe opisanie jazyka i sistemnaja leksikografija* [Chosen Works. Integrated description of language and systematic lexicography], Moscow.
- Apresjan, J.D. 1974, *Leksičeskaja semantika : sinonimičeskie sredstva jazyka*, Nauka, Moscow.
- Archangel'skij, V.L. 1964, *Ustojčivye frazy v sovremennom russkom jazyke*, Izdatel'stvo Rostovskogo universiteta, Rostov-na-Donu.
- Arppe, A. 2008, *Univariate, bivariate, and multivariate methods in corpus-based lexicography : a study of synonymy*, University of Helsinki, Helsinki.
- Avdeeva, O.I. 2004, *Vnešnjaja i vnutrennjaja sintagmatika frazeologičeskikh edinic : učebnoe posobie k speckursu*, Izdatel'stvo Adygejskogo universiteta, Majkop.
- Backhouse, A.E. 1992, "Connotation" in *International encyclopedia of linguistics*, ed. W. Bright, Oxford University Press, New York NY, pp. 297-298.
- Baranov, A. & Dobrovol'skij, D. 1999, "Idioms from a cognitive perspective", *Moscow State University Bulletin* 19, vol. 1, pp. 64-75.
- Baranov, A. & Dobrovol'skij, D. 1996, "Cognitive modeling of actual meaning in the field of phraseology", *Journal of Pragmatics*, vol. 25, no. 3, pp. 409-429.
- Baranov, A. & Dobrovol'skij, D. 1990, "Struktury znaniy i ih jazykovaja ontologizacija v znachenii idiomy [Structures of knowledge and their linguistic ontologization in idiom meaning]", *Učenie zapiski Tartuskogo universiteta* [Scientific papers from Tartu university], vol. 903, pp. 20-37.
- Bar-Hillel, Y. 1955, "Idioms" in *Machine translation of languages*, eds. W.N. Locke & A.D. Booth, MIT Press, Cambridge, MA, pp. 183-193.
- Bar-Hillel, Y. 1953, "Some linguistic problems connected with machine translation", *Philosophy of Science*, no. 20, pp. 217-225.
- Barkema, H. 1996, "Idiomaticity and terminology : a multi-dimensional descriptive model", *Studia Linguistica*, vol. 50, no. 2, pp. 125-160.

- Barthes, R. 1972, *Mythologies*, Hill and Wang, New York.
- Bartmiński, J. (ed) 1980, *Słownik ludowych stereotypów językowych. Zeszyt próbny [Encyclopedia of Folk Stereotypes. First Draft]*, Wydawnictwa Uniwersytetu Wrocławskiego, Wrocław.
- Bobrow, S.A. & Bell, S.M. 1973, "On catching on to idiomatic expressions", *Memory and Cognition*, vol. 1, pp. 343-346.
- Bojčenko, V.V. 1993, *Individual'no-avtorskie preobrazovanija frazeologičeskikh edinic*, Candidate of philological sciences thesis, St. Petersburg state university, St. Petersburg.
- Bolinger, D. 1968, "Entailment and the meaning of structures", *Glossa*, vol. 2, no. 2, pp. 119-127.
- Bresnan, J. 1982, *The mental representation of grammatical relations*, The MIT Press, Cambridge, MA.
- Burger, H. 1998, *Phraseologie : eine Einführung am Beispiel des Deutschen*, Erich Schmidt, Berlin.
- Burger, H. 1973, *Idiomatik des Deutschen*, Niemeyer, Tübingen.
- Burger, H., Buhofer, A. & Sialm, A. 1982, *Handbuch der Phraseologie*, de Gruyter, Berlin.
- Cacciari, C. 1993, "The place of idioms in a literal and metaphorical world" in *Idioms : processing, structure, and interpretation*, eds. C. Cacciari & P. Tabossi, Lawrence Erlbaum Associates, Hillsdale, pp. 27-56.
- Cacciari, C., Corradini, P. & Padovani, R. 2005, "Speed of processing effects on spoken idioms comprehension", *Proceedings of the XXVII Annual Meeting of the Cognitive Science Society*.
- Cacciari, C. & Glucksberg, S. 1994, "Understanding figurative language" in *Handbook of Psycholinguistics*, ed. M.A. Gernsbacher, Academic Press, pp. 447-477.
- Cacciari, C. & Glucksberg, S. 1991, "Understanding idiomatic expressions : the contribution of word meanings" in *Understanding word and sentence*, ed. G.B. Simpson, Elsevier Science Publishers, North-Holland, pp. 217-240.
- Cacciari, C. & Tabossi, P. 1988, "The comprehension of idioms", *Journal of Memory and Language*, vol. 27, pp. 668-683.
- Carston, R. 2002, *Thoughts and utterances : the pragmatics of explicit communication*, Blackwell, Oxford.
- Cavell, S. 1976, *Must we mean what we say : a book of essays*, Cambridge university press, Cambridge.
- Čermák, F. 2007, "Idioms and morphology" in *Phraseologie / Phraseology Ein internationales Handbuch zeitgenössischer Forschung / An International Handbook of Contemporary Research*, eds. H. Burger, D. Dobrovol'skij, P. Kühn & N.R. Norick, Mouton De Gruyter, Berlin/New York, pp. 20-26.
- Chafe, W.L. 1968, "Idiomaticity as an anomaly in the Chomskyan paradigm", *Foundations of Language*, vol. 4, pp. 109-127.
- Chandler, D. 2002, *Semiotics : the basics*, Routledge, London.

- Chomsky, N. 1970, "Remarks on nominalizations" in *Readings in English transformational grammar*, eds. R.A. Jacobs & P.S. Rosenbaum, Georgetown University Press, Washington, D.C., pp. 184-221.
- Chomsky, N. 1980, *Rules and representations*, Columbia University Press, New York.
- Chomsky, N. 1975, *Reflections on language*, Pantheon, New York.
- Chomsky, N. 1965, *Aspects of the theory of syntax*, MIT Press, Cambridge MA.
- Chomsky, N. 1957, *Syntactic structures*, Mouton, The Hague.
- Clark, E.V. 1987, "The principle of contrast : a constraint on language acquisition" in *Mechanisms of language acquisition*, ed. B. MacWhinney, Lawrence Erlbaum, Hillsdale, N.J., pp. 1-33.
- Clark, H.H. & Lucy, P. 1975, "Understanding what is meant from what is said : a study in conversationally conveyed requests", *Journal of Verbal Learning and Verbal Behaviour*, vol. 14, no. 1, pp. 56-72.
- Coates, J. 1990, "Modal meaning : the semantic-pragmatic interface", *Journal of Semantics*, vol. 7, pp. 53-63.
- Colombo, L. 1998, "Role of context in the comprehension of ambiguous Italian idioms" in *Sentence processing : a crosslinguistic perspective*, ed. J. Runner, Emerald Group Publishing Limited, pp. 379-404.
- Colombo, L. 1993, "The comprehension of ambiguous idioms in context" in *Idioms : processing, structure, and interpretation*, eds. C. Cacciari & P. Tabossi, Lawrence Erlbaum Associates, Hillsdale, pp. 163-200.
- Cook, G. 1992, *The discourse of advertising*, Routledge, London.
- Cowie, A.P. (ed) 1998, *Phraseology : theory, analysis, and applications*, Clarendon Press, Oxford.
- Cowie, A.P. 1981, "The treatment of collocations and idioms in learner's dictionaries", *Applied Linguistics*, vol. 2, no. 3, pp. 223-235.
- Cowie, A.P. & Mackin, R. 1984, *Oxford dictionary of current idiomatic English, volume 1. Verbs with prepositions and particles*, Oxford University Press, Oxford.
- Cowie, A.P., Mackin, R. & McCaig, I.R. 1993, *Oxford dictionary of English idioms*, Oxford University Press, Oxford.
- Cowie, A.P., Mackin, R. & McCaig, I.R. 1983, *Oxford dictionary of current idiomatic English, volume 2. Phrase, clause and sentence idioms*, Oxford University Press, Oxford.
- Croft, W. 2001, *Radical construction grammar : syntactic theory in typological perspective*, Oxford University Press, Oxford.
- Croft, W. & Cruse, D.A. 2004, *Cognitive linguistics*, Cambridge University Press, Cambridge.
- Crystal, D. 1997, *A dictionary of linguistics and phonetics*, 4th edn, Blackwell, Oxford.
- Culicover, P.W. & Jackendoff, R. 2005, *Simpler syntax*, Oxford University Press, Oxford.
- Culicover, P. & Wilkins, W. 1986, "Control, PRO, and the projection principle", *Language*, vol. 62, pp. 112-153.

- Cutting, J.C. & Bock, K. 1997, "That's the way the cookie bounces : syntactic and semantic components of experimentally elicited idiom blends", *Memory & Cognition*, vol. 25, no. 1, pp. 57-71.
- Dal', V.I. 2000, *Tolkovyj slovar' russkogo âzyka : sovremennââ versiâ*, Èkmo-Press, Moskva.
- Davies, M. 1982, "Idiom and metaphor", *Proceedings of the Aristotelian Society*, no. 83, pp. 67-85.
- de Haan, F. 1997, *The interaction of modality and negation : a typological study*, Garland Pub., New York.
- Devos, F. 2003, "Semantic vagueness and lexical polyvalence", *Studia Linguistica*, vol. 57, no. 3, pp. 121-141.
- Dobrovol'skij, D. 1998, "On cultural component in the semantic structure of idioms", *Europhras '97. Phraseology and Paremiology*, ed. P. Ďurčo, Akadémia PZ, Bratislava, pp. 55-61.
- Dobrovol'skij, D. 1995, *Kognitive Aspekte der Idiom-Semantik : Studien zum Thesaurus deutscher Idiome*, Narr, Tübingen.
- Dobrovol'skij, D. 1992, "Phraseological universals : theoretical and applied aspects" in *Meaning and grammar : cross-linguistic perspectives*, eds. M. Kefer & J. van der Auwera, Mouton de Gruyter, Berlin ; New York, pp. 279-301.
- Dobrovol'skij, D. 1988, *Phraseologie als Objekt der Universalienlinguistik*, Enzyklopädie, Leipzig.
- Dobrovol'skij, D. & Piirainen, E. 2005, *Figurative language : cross-cultural and cross-linguistic perspectives*, Elsevier, Oxford.
- Dobrydneva, E.A. 2000, *Kommunikativno-pragmaticseskaja paradigma russkoj frazeologii*, Peremena, Volgograd.
- Dong, Q.P. 1971, "The applicability of transformations to idioms", *Papers from the Seventh Regional Meeting*, Chicago Linguistic Society, Chicago, pp. 198-205.
- Edinburgh Associative Thesaurus (EAT) 2011, **Edinburgh Associative Thesaurus (EAT)** [Internet], Science and Technology Facilities Council e-Science. Available from: <<http://www.eat.rl.ac.uk>> [Accessed 17 May 2009]
- Estill, R.B. & Kemper, S. 1982, "Interpreting idioms", *Journal of Psycholinguistic Research*, vol. 11, pp. 559-568.
- Everaert, M. & Kuiper, K. 1996, "Theory and data in idiom research", *Papers from the Parasession on Theory and Data in Linguistics*, eds. M. AuCoin, R. Celis, L.M. Dobrin, L. McNair & K. Singer, Chicago Linguistic Society, Chicago, pp. 43-57.
- Fauconnier, G. 1997, *Mappings in thought and language*, Cambridge U.P., Cambridge.
- Fauconnier, G. 1994, *Mental spaces : aspects of meaning construction in natural language*, Cambridge University Press, Cambridge.
- Fauconnier, G. & Turner, M. 2002, *The way we think : conceptual blending and the mind's hidden complexities*, Basic Books, New York.
- Fauconnier, G. & Turner, M. 1998, "Principles of conceptual integration" in *Discourse and cognition : bridging the gap*, ed. J. Koenig, Center for the Study of Language and Information, Stanford CA, pp. 269-283.

- Fauconnier, G. & Turner, M. 1996, "Blending as a central process of grammar" in *Conceptual structure, discourse and language*, ed. A.E. Goldberg, Center for the Study of Language and Information, Stanford CA, pp. 113-130.
- Fellbaum, C. (ed) 2007, *Idioms and collocations : corpus-based linguistic and lexicographic studies*, Continuum, London.
- Fernando, C. 1996, *Idioms and idiomaticity*, Oxford University Press, Oxford.
- Fillmore, C.J. 1999, "Inversion and constructional inheritance" in *Lexical and constructional aspects of linguistic explanation*, eds. G. Webelhuth, J. Koenig & A. Kathol, Center for the Study of Language and Information, Stanford, CA, pp. 113-128.
- Fillmore, C.J. 1968, "The case for case" in *Universals in linguistic theory*, eds. E. Bach & R.T. Harms, Holt, Rinehart & Winston, New York, pp. 1-88.
- Fillmore, C.J., Johnson, C.R. & Petruck, M.R.L. 2003, "Background to Framenet", *International Journal of Lexicography*, vol. 16, no. 3, pp. 235-250.
- Fillmore, C.J. & Kay, P. 1996, *Construction grammar*, Unpublished ms, University of California at Berkeley Department of linguistics.
- Fillmore, C.J. & Kay, P. 1993, *Construction grammar coursebook*, Unpublished ms, University of California at Berkeley Department of linguistics.
- Fillmore, C.J., Kay, P. & O'Connor, M.C. 1988, "Regularity and idiomaticity in grammatical constructions : the case of *let alone*", *Language*, vol. 64, no. 3, pp. 501-538.
- Fillmore, C.J., Petruck, M.R.L., Ruppenhofer, J. & Wright, A. 2003, "Framenet in action : the case of attaching", *International Journal of Lexicography*, vol. 16, no. 3, pp. 297-332.
- Flores d'Arcais, G.B. 1993, "The comprehension and semantic interpretation of idioms" in *Idioms : processing, structure, and interpretation*, eds. C. Cacciari & P. Tabossi, Lawrence Erlbaum Associates, Hillsdale New Jersey, pp. 79-98.
- Fodor, J.A. 1983, *The modularity of mind : an essay on faculty psychology*, MIT Press, Cambridge, MA.
- Fodor, J., Garrett, M., Walker, E. & Parkes, C. 1980, "Against definitions", *Cognition*, vol. 8, pp. 263-367.
- Fontenelle, T. 1998, "Discovering significant lexical functions in dictionary entries" in *Phraseology : theory, analysis, and applications*, ed. A.P. Cowie, Clarendon Press, Oxford, pp. 189-207.
- Fraser, B. 1970, "Idioms within a transformational grammar", *Foundations of Language*, vol. 6, pp. 22-42.
- Fried, M., Boas, H.C., Leino, J. & Östman, J. 2005, *Grammatical constructions : back to the roots*, John Benjamins Publishing, Amsterdam.
- Fried, M. & Östman, J. 2004, "Construction grammar : a thumbnail sketch" in *Construction grammar in a cross-language perspective*, eds. M. Fried & J. Östman, John Benjamins Publishing Company, Amsterdam, pp. 11-86.
- Gazdar, G., Klein, E., Pullum, G.K. & Sag, I. 1985, *Generalized phrase structure grammar*, Blackwell, Oxford.



- Geeraerts, D. 1995, "Specialization and reinterpretation on idioms" in *Idioms : structural and psychological perspectives*, eds. M. Everaert, E. van der Linden, A.Y. Schenk & R. Schreuder, Lawrence Erlbaum Associates, Hillsdale New Jersey, pp. 57-74.
- Gernsbacher, M.A. & Robertson, R.R.W. 1999, "The role of suppression in figurative language comprehension", *Journal of Pragmatics*, vol. 31, pp. 1616-1630.
- Gibbs, R.W., Jr 2001, "Evaluating contemporary models of figurative language understanding", *Metaphor and Symbol*, vol. 16, no. 3-4, pp. 317-333.
- Gibbs, R.W., Jr 2000, "Making good psychology out of blending theory", *Cognitive Linguistics*, vol. 11, pp. 347-358.
- Gibbs, R.W., Jr 1994, *The poetics of mind : figurative thought, language, and understanding*, Cambridge U.P., Cambridge.
- Gibbs, R.W., Jr 1993, "Why idioms are not dead metaphors" in *Idioms : processing, structure, and interpretation*, eds. C. Cacciari & P. Tabossi, Lawrence Erlbaum Associates, Hillsdale New Jersey, pp. 57-77.
- Gibbs, R.W., Jr 1990, "Psycholinguistic studies on the conceptual basis of idiomaticity", *Cognitive Linguistics*, vol. 1, pp. 417-451.
- Gibbs, R.W., Jr 1986, "Skating on thin ice : literal meaning and understanding idioms in conversation", *Discourse Processes*, vol. 7, pp. 17-30.
- Gibbs, R.W., Jr 1985, "On the process of understanding idioms", *Journal of Psycholinguistic Research*, vol. 14, pp. 465-472.
- Gibbs, R.W., Jr 1980, "Spilling the beans on understanding and memory for idioms in conversation", *Memory and Cognition*, vol. 8, pp. 149-156.
- Gibbs, R.W., Jr & Nayak, N.P. 1989, "Psycholinguistic studies on the syntactic behaviour of idioms", *Cognitive Psychology*, vol. 21, pp. 100-138.
- Gibbs, R.W., Jr, Nayak, N.P., Bolton, J.L. & Keppel, M.E. 1989, "Speakers' assumptions about the lexical flexibility of idioms", *Memory and Cognition*, vol. 17, pp. 58-68.
- Gibbs, R.W., Jr, Nayak, N.P. & Cutting, J.C. 1989, "How to kick the bucket and not decompose : analyzability and idiom processing", *Journal of Memory and Language*, vol. 28, pp. 567-593.
- Gläser, R. 1988, "The grading of idiomaticity as a presupposition for a taxonomy of idioms" in *Understanding the lexicon : meaning, sense and world knowledge in lexical semantics*, eds. W. Hüllen & R. Schulze, Max Niemeyer Verlag, Tübingen, pp. 264-277.
- Glucksberg, S. 2001, *Understanding figurative language : from metaphors to idioms*, Oxford University Press, Oxford.
- Glucksberg, S. & Keysar, B. 1993, "How metaphors work" in *Metaphor and Thought*, ed. A. Ortony, 2nd edn, Cambridge University Press, Cambridge, pp. 401-424.
- Goldberg, A.E. 2006, *Constructions at work : the nature of generalization in language*, Oxford University Press, Oxford.
- Goldberg, A.E. 1995, *Constructions : a construction grammar approach to argument structure*, University of Chicago press, Chicago.
- Goldberg, A.E. & Jackendoff, R. 2004, "The English resultative as a family of constructions", *Language*, vol. 80, pp. 532-568.

- Goldsmith, J.A. 1990, *Autosegmental and metrical phonology*, Blackwell, Oxford.
- Goldsmith, J.A. 1979, *Autosegmental phonology*, Garland, New York.
- Google Groups 2011, **Google Groups** [Internet], Google Inc. Available from: <<http://groups.google.com>> [Accessed 15 Jan 2007]
- Grice, H.P. 1975, "Logic and conversation" in *Syntax and Semantics*, eds. P. Cole & J.L. Morgan, Academic Press, New York, pp. 41-58.
- Gruber, J.S. 1965, *Studies in lexical relations*, Ph.D. thesis, MIT Cambridge, Mass.
- Haarala, R. 1990, *Suomen kielen perussanakirja*, Valtion painatuskeskus, Helsinki.
- Haiman, J. 1985, "Symmetry" in *Iconicity in syntax : proceedings of a symposium on iconicity in syntax, Stanford, June 24-6, 1983*, ed. J. Haiman, John Benjamins, Amsterdam, pp. 73-95.
- Häkkinen, K. 2000, "Idiomin määritelmää" in *Fennistica fausta in honorem Mauno Koski septuagenarii = Juhlakirja Mauno Kosken 70-vuotispäivän kunniaksi 21.2.2000*, eds. E. Lähdemäki & Y. Bertills, Åbo Akademi, Åbo, pp. 1-15.
- Hakulinen, A. & Karlsson, F. 1988, *Nykysuomen lauseoppia*, 2nd edn, Suomalaisen kirjallisuuden seura, Helsinki.
- Hakulinen, A., Karlsson, F. & Vilks, M. 1980, *Suomen tekstilauseiden piirteitä : kvantitatiivinen tutkimus*, University of Helsinki, Helsinki.
- Hakulinen, A., Vilks, M., Korhonen, R., Koivisto, V., Heinonen, T.R. & Alho, I. 2004, *Iso suomen kielioppi*, Suomalaisen Kirjallisuuden Seura, Helsinki.
- Halliday, M.A.K. 1985, *An introduction to functional grammar*, Arnold, London.
- Halliday, M.A.K. & Hasan, R. 1976, *Cohesion in English*, Longman, London.
- Harder, P. 2003, "Mental spaces : exactly when do we need them?", *Cognitive Linguistics*, vol. 14, pp. 91-96.
- Healey, A. 1968, "English idioms", *KIVUNG, (Journal of the Linguistic Society of the University of Papua New Guinea)*, vol. 1, no. 2, pp. 71-108.
- Heinonen, T.R. 2007, "Variation and flexibility within verb idioms in Finnish", *Collocations and idioms 1 : Papers from the First Nordic Conference on Syntactic Freezes, Joensuu, May 19-20, 2006*, eds. M. Nenonen & S. Niemi, University of Joensuu, Joensuu, pp. 146-159.
- Herold, A. 2007, "Corpus queries" in *Idioms and collocations : corpus-based linguistic and lexicographic studies*, ed. C. Fellbaum, Continuum, London, pp. 54-63.
- Hewings, A. & Hewings, M. 2005, *Grammar and context : an advanced resource book*, Routledge ; New York, London.
- Hockett, C.F. 1958, *A course in modern linguistics*, Macmillan, New York.
- Hoffmann, S. 2007, "Processing Internet-derived text – creating a corpus of Usenet messages", *Literary and Linguistic Computing*, vol. 22, no. 2, pp. 151-165.
- Holmberg, A. & Nikanne, U. 1993, *Case and other functional categories in Finnish syntax*, Mouton de Gruyter, Berlin.
- Howarth, P.A. 1996, *Phraseology in English academic writing : some implications for language learning and dictionary making*, Max Niemeyer, Tübingen.
- Hyvärinen, I. 2007, "Phraseologie des Finnischen" in *Phraseologie / Phraseology Ein internationales Handbuch zeitgenössischer Forschung / An International Handbook of*

- Contemporary Research*, eds. H. Burger, D. Dobrovol'skij, P. Kühn & N.R. Norick, Mouton de Gruyter, Berlin, pp. 737-751.
- Hyvärinen, I. 2004, "Kausative Resultativkonstruktionen als Phraseoschablonen im Deutschen und im Finnischen", *Phraseologismen als Gegenstand sprach- und kulturwissenschaftlicher Forschung. Akten der Europäischen Gesellschaft für Phraseologie (EUROPHRAS) und des Westfälischen Arbeitskreises "Phraseologie/ Parömio-logie"* (Loccum 2002), eds. C. Földes & J. Wirrer, Schneider Verlag, Baltmannsweiler, pp. 135-146.
- Hyvärinen, I. 1996a, "Zur Passivierbarkeit von finnischen Verbididiomen" in *Studien zur Phraseologie des Deutschen und des Finnischen II.*, ed. J. Korhonen, Brockmeyer, Bochum, pp. 53-68.
- Hyvärinen, I. 1996b, "Zur Semantik von deutschen und finnischen Verbididiomen" in *Studien zur Phraseologie des Deutschen und des Finnischen II.*, ed. J. Korhonen, Brockmeyer, Bochum, pp. 345-439.
- Hyvärinen, I. 1996c, "Zur syntaktischen Negierbarkeit von finnischen Verbididiomen" in *Studien zur Phraseologie des Deutschen und des Finnischen II*, ed. J. Korhonen, Brockmeyer, Bochum, pp. 81-98.
- Hyvärinen, I. 1992a, "Zum Festigkeitsgrad des Verbs und zur aktionalen Reihenbildung bei deutschen und finnischen Verbididiomen. Ein Werkstattbericht im Rahmen des Projekts „Kontrastive Verbididiomatik Deutsch – Finnisch“" in *Phraseologie und Wortbildung – Aspekte der Lexikonerweiterung*, ed. J. Korhonen, Max Niemeyer Verlag, Tübingen, pp. 21-38.
- Hyvärinen, I. 1992b, "Zur Antonymie in deutsch-finnischer Verbididiomatik", *Untersuchungen zur Phraseologie des Deutschen und anderer Sprachen : einzelsprachspezifisch – kontrastiv – vergleichend. Internationale Tagung in Turku 6.–7.9.1991*, ed. J. Korhonen, Lang, Frankfurt am Main, pp. 79-104.
- Ingo, R. 2000, "Idiomi toisen kielen semanttis-pragmaattisena käännösvastineena" in *Fennistica fausta in honorem Mauno Koski septuagenarii = Juhlakirja Mauno Kosken 70-vuotispäivän kunniaksi 21.2.2000*, eds. E. Lähdemäki & Y. Bertills, Åbo Akademi, Åbo, pp. 33-45.
- Iordanskaja, L.N. & Mel'čuk, I.A. 1980, "Konnotacija v lingvističeskoj semantike", *Wiener Slawistischer Almanach*, vol. 6, pp. 191-210.
- Itkonen, E. 1966, *Kieli ja sen tutkimus*, Wsoy, Porvoo.
- Jackendoff, R. 2007a, *Language, consciousness, culture : essays on mental structure*, MIT Press, Cambridge MA.
- Jackendoff, R. 2007b, "Linguistics in cognitive science : the state of the art", *The Linguistic Review*, vol. 24, no. 4, pp. 347-401.
- Jackendoff, R. 2007c, "A parallel architecture perspective on language processing", *Brain Research*, vol. 1146, pp. 2-22.
- Jackendoff, R. 2002a, "English particle constructions, the lexicon, and the autonomy of syntax" in *Verb-Particle Explorations*, eds. N. Dehé, R. Jackendoff, A. McIntyre & S. Urban, Mouton de Gruyter, Berlin, pp. 67-94.
- Jackendoff, R. 2002b, *Foundations of language : brain, meaning, grammar, evolution*, Oxford University Press, Oxford.

- Jackendoff, R. 2002c, "The mind doesn't work that way : the scope and limits of computational psychology (review)", *Language*, vol. 78, no. 1, pp. 164-170.
- Jackendoff, R. 1997, *The architecture of the language faculty*, MIT Press, Cambridge MA.
- Jackendoff, R. 1996a, "The architecture of the linguistic-spatial interface" in *Language and Space*, eds. P. Bloom, M.A. Peterson, L. Nadel & M.F. Garrett, MIT Press, Cambridge MA, pp. 1-30.
- Jackendoff, R. 1996b, "How language helps us think", *Pragmatics and Cognition*, vol. 4, pp. 1-24.
- Jackendoff, R. 1995, "The boundaries of the lexicon" in *Idioms : structural and psychological perspectives*, eds. M. Everaert, E. van der Linden, A. Schenk & R. Schreuder, Lawrence Erlbaum Associates, Hillsdale New Jersey, pp. 133-165.
- Jackendoff, R. 1994, *Patterns in the mind*, Basic Books, New York.
- Jackendoff, R. 1993, *Languages of the mind : essays on mental representation*, MIT Press, Cambridge MA.
- Jackendoff, R. 1990, *Semantic structures*, MIT Press, Cambridge MA.
- Jackendoff, R. 1988, "Why are they saying these things about us? (Topic-Comment column)", *Natural Language and Linguistic Theory*, vol. 6, pp. 435-442.
- Jackendoff, R. 1987a, *Consciousness and the computational mind*, MIT Press, Cambridge MA.
- Jackendoff, R. 1987b, "The status of thematic relations in linguistic theory", *Linguistic Inquiry*, vol. 18, pp. 369-411.
- Jackendoff, R. 1983, *Semantics and cognition*, MIT Press, Cambridge MA.
- Jackendoff, R. 1976, "Towards an explanatory semantic representation", *Linguistic Inquiry*, vol. 7, pp. 89-150.
- Jackendoff, R. 1972, *Semantic interpretation in generative grammar*, MIT Press, Cambridge MA.
- Jackendoff, R. & Aaron, D. 1991, "Review of more than cool reason : a field guide to poetic metaphor by George Lakoff and Mark Turner", *Language*, vol. 67, pp. 320-388.
- Jacobs, R.A., Rosenbaum, P.S. & Postal, P.M. 1968, *English transformational grammar*, Blaisdell Publishing Company, Waltham MA.
- Janus, R.A. & Bever, T.G. 1985, "Processing of metaphoric language : an investigation of the 3-stage model of metaphor comprehension", *Journal of Psycholinguistic Research*, vol. 14, no. 5, pp. 473-489.
- Järviö-Nieminen, I. 1959, *Suomalaiset sanomukset*, Suomalaisen Kirjallisuuden Seura, Helsinki.
- Jespersen, O. 1924, *The philosophy of grammar*, Allen & Unwin, London.
- Karaulov, Ū.N. 1996, *Russkij associativnyj slovar' [Russian associative dictionary]*, vol. 3, IrÂ Ran, Moscow.
- Karaulov, Ū.N. 1994a, *Russkij associativnyj slovar' [Russian associative dictionary]*, vol. 2, IrÂ Ran, Moscow.
- Karaulov, Ū.N. 1994b, *Russkij associativnyj slovar' [Russian associative dictionary]*, vol. 1, Pomovskij i partnery, Moscow.

- Kari, E. 1993, *Naulan kantaan. Nykysuomen idiomisanakirja*, Otava, Helsingissä.
- Karlsson, F. 1983, *Suomen kielen äänne- ja muotorakenne*, Söderström, Porvoo.
- Karlsson, G. 2000, "Lehtiväen idiomissa erikoinen monikon käyttötapa" in *Fennistica fausta in honorem Mauno Koski septuagenarii = Juhlakirja Mauno Kosken 70-vuotispäivän kunniaksi 21.2.2000*, eds. E. Lähdemäki & Y. Bertills, Åbo Akademi, Åbo, pp. 96-104.
- Katz, J.J. & Fodor, J. 1963, "The structure of semantic theory", *Language*, vol. 39, pp. 170-210.
- Katz, J.J. & Postal, P.M. 1963, "Semantic interpretation of idioms and sentences containing them", *M.I.T Research Laboratory of Electronics, Quarterly Progress Report*, vol. 70, pp. 275-282.
- Kay, P. 1995, "Construction grammar" in *Handbook of pragmatics : manual*, eds. J. Verschueren, J. Östman & J. Blommaert, John Benjamins Publishing Company, Amsterdam, pp. 171-177.
- Kay, P. 1984, "The kinda/sorta construction", *Proceedings of the Tenth Annual Meeting of the Berkeley Linguistics Society*, eds. C. Brugman & M. Macauley, University of California, Berkeley CA, pp. 157-171.
- Kay, P. & Fillmore, C.J. 1999, "Grammatical constructions and linguistic generalizations : the *What's X doing Y?* construction", *Language*, vol. 75, pp. 1-33.
- Kempson, R.M. 1977, *Semantic theory*, Cambridge University Press, Cambridge.
- Keysar, B. & Bly, B. 1995, "Intuitions of the transparency of idioms : can one keep a secret by spilling the beans?", *Journal of Memory and Language*, vol. 34, pp. 89-109.
- Keysar, B. & Bly, B.M. 1999, "Swimming against the current : do idioms reflect conceptual structure?", *Journal of Pragmatics*, vol. 31, pp. 1559-1578.
- Kielikone Ltd. 2010, *MOT Englanti 4.8 sanakirja*, Helsinki.
- Kielipankki 2011, **Kielipankki** [Internet], CSC Finnish IT center for science, Helsinki. Available from: <<http://www.csc.fi/tutkimus/alat/kielitiede>> [Accessed 8 Feb 2009]
- Kittredge, R., Iordanskaja, L. & Polguère, A. 1988, "Multi-lingual text generation and the Meaning-Text Theory", *Proceedings of the 2nd International Conference on Theoretical and Methodological Issues in Machine Translation of Natural Languages*, Carnegie Mellon University, Pittsburgh, PA.
- Kiuru, S. 1977, *Suomen kielen kieltohakuiset verbit : murreaineistoon perustuva syntaktis-semanttinen tutkimus*, SKS, Helsinki.
- Kivimies, Y. 1964, *Näinkin voi sanoa : suomen kielen fraseologiaa*, Tammi, Helsinki.
- Kobozeva, I.M. 2000, *Lingvističeskaâ semantika*, Èditorial URSS, Moscow.
- Kooij, J.G. 1971, *Ambiguity in natural language : an investigation of certain problems in its linguistic description*, North-Holland Pub. Co, Amsterdam.
- Korhonen, J. 2000, "Zum Einfluss des Hochdeutschen auf die Lexik und Phraseologie des Finnischen", *Hochdeutsch in Skandinavien : internationales Symposium, Zürich, 14.-16. Mai 1998*, eds. H. Naumann & S. Müller, A. Francke, Tübingen, pp. 99-115.
- Korhonen, J. 1998, "Zur Übersetzung von Verbidiomen in Kriminalserien. Am Beispiel Deutsch und Finnisch" in *Europäische Phraseologie im Vergleich : gemeinsames Erbe und kulturelle Vielfalt*, ed. W. Eisman, Universitätsverlag Brockmeyer, Bochum, pp. 415-430.

- Korhonen, J. 1996, *Studien zur Phraseologie des Deutschen und des Finnischen II*. Universitätsverlag Brockmeyer, Bochum.
- Korhonen, J. 1995, *Studien zur Phraseologie des Deutschen und des Finnischen*, Universitätsverlag Brockmeyer, Bochum.
- Korhonen, J. 1991, "Kontrastive Verbidiomatik Deutsch-Finnisch. Ein Forschungsbericht" in *Sprichwörter und Redensarten im interkulturellen Vergleich*, eds. A. Sabban & J. Wirrer, Westdeutscher Verlag, Opladen, pp. 37-65.
- Korhonen, J. 1987, "Überlegungen zum Forschungsprojekt „Kontrastive Verbidiomatik Deutsch-Finnisch“, *Beiträge zur allgemeinen und germanistischen Phraseologieforschung : internationales Symposium in Oulu 13.-15. Juni 1986*, ed. J. Korhonen, Oulun yliopisto, Oulu, pp. 1-22.
- Korhonen, R. 1993, *Buts about conjunctions : a syntactic study of conjunction expressions in Finnish*, Suomalaisen Kirjallisuuden Seura, Helsinki.
- Korhonen, R. & Vilkuna, M. 2005, **Sananselityksiä : Ison suomen kieliopin termejä** [Internet], Kotimaisten kielten tutkimuskeskus, Helsinki. Available from: <<http://kaino.kotus.fi/www/verkkojulkaisut/julk1>> [Accessed 13 March 2007].
- Koski, M. 1981, "Mitä leksikaalistuminen on?", *Sananmuodostuksen ongelmia, SKYJ*, vol. 7, pp. 5-28.
- Kreidler, C.W. 1998, *Introducing English semantics*, Routledge, London.
- Krongauz, M.A. 2001, *Semantika*, Rossijskij gumanitarnyj universitet, Moscow.
- Kunin, A.V. 1996, *Frazeologija sovremennogo anglijskogo jazyka*, Vysšaja škola, Moscow.
- Kunin, A.V. 1964, *Osnovnye ponjatija anglijskoj frazeologii kak lingvističeskoj discipliny i sozdanie anglo-russkogo frazeologičeskogo slovarja*, Doctor of philological sciences thesis, Maurice Thorez Moscow State Pedagogical Institute of Foreign Languages, Moscow.
- Kunin, A.V. 1956, *Anglo-russkij frazeologičeskij slovar'* [English-Russian phraseological dictionary], 2nd edn, Russkij jazyk, Moscow.
- Kurikka, U. 1979, *Kielto 1960-luvun suomen yleiskielessä = Negation in the standard Finnish of the 1960's*, Oulun yliopisto, Oulu.
- Kuusi, A. 1971, *Johdatusta suomen kielen fraseologiaan*, Suomalaisen Kirjallisuuden Seura, Helsinki.
- Kuusi, M. 1963, "Sananparsiston rakenneanalyysin terminologiaa", *Virittäjä*, vol. 67, pp. 339-347.
- Kuusi, M. 1960, *Suomen kansan vertauksia*, Suomalaisen Kirjallisuuden Seura, Helsinki.
- Kuusi, M. 1952, "Nimityksistä sananparsi, sananlasku, puheenparsi ja lauseparsi", *Virittäjä*, vol. 56, pp. 66-68.
- Lagerspetz, O. 2006, *Smuts : en bok om världen, vårt hem*, Brutus Östlings bokförlag Symposion, Eslöv.
- Lähdemäki, E. & Bertills, Y. (eds) 2000, *Fennistica fausta in honorem Mauno Koski septuagenarii = Juhlakirja Mauno Kosken 70-vuotispäivän kunniaksi 21.2.2000*, Åbo Akademi, Åbo.
- Laitinen, L. 1993, "Nesessiivirakenne, kieliopillistuminen ja subjektiivisuus", *Virittäjä*, vol. 97, pp. 149-170.

- Lakoff, G. 1993, "The contemporary theory of metaphor" in *Metaphor and Thought*, ed. A. Ortony, 2nd edn, Cambridge University Press, Cambridge, pp. 202-251.
- Lakoff, G. 1987, *Women, fire, and dangerous things : what categories reveal about the mind*, University of Chicago Press, Chicago.
- Lakoff, G. 1986, "The meanings of literal", *Metaphor and Symbolic Activity*, vol. 1, pp. 291-296.
- Lakoff, G. 1970, "A note on ambiguity and vagueness", *Linguistic Inquiry*, vol. 1, pp. 357-359.
- Lakoff, G. & Johnson, M. 1980, *Metaphors we live by*, University of Chicago Press, Chicago IL.
- Lakoff, G. & Turner, M. 1989, *More than cool reason : a field guide to poetic metaphor*, University of Chicago Press, Chicago.
- Langacker, R.W. 1991, *Foundations of cognitive grammar. Vol. 2, Descriptive application*, Stanford University Press, Stanford, CA.
- Langacker, R.W. 1987, *Foundations of cognitive grammar. Vol. 1, Theoretical prerequisites*, Stanford University Press, Stanford, CA.
- Langlotz, A. 2006, *Idiomatic creativity : a cognitive-linguistic model of idiom-representation and idiom-variation in English*, John Benjamins Publishing Company, Philadelphia.
- Larjavaara, M. 2007, *Pragmasemantiikka*, Suomalaisen Kirjallisuuden Seura, Helsinki.
- Larjavaara, M. 1990, *Suomen deiksis*, Suomalaisen Kirjallisuuden Seura, Helsinki.
- Laukkanen, K., Hakamies, P., Kuusi, M. & Vuori, P. 1978, *Sananlaskut*, Suomalaisen Kirjallisuuden Seura, Helsinki.
- LDCE = *Longman dictionary of contemporary English*, see Summers, D. 1987.
- Leezenberg, M. 2001, *Contexts of metaphor*, Elsevier, Amsterdam.
- Leino, J. & Östman, J. 2005, "Constructions and variability" in *Grammatical constructions : back to the roots*, eds. M. Fried, H.C. Boas, J. Leino & J. Östman, John Benjamins Publishing, Amsterdam, pp. 191-213.
- Leino, P., Helasvuo, M.-L., Lauerma, P., Nikanne, U. & Onikki, T. (eds) 1990, *Suomen kielen paikallissijat konseptuaalisessa semantiikassa*, Helsingin yliopisto, Helsinki.
- Lieberman, M. & Prince, A. 1977, "On stress and linguistic rhythm", *Linguistic Inquiry*, vol. 8, pp. 249-336.
- Lubensky, S. 1997, *Russko-anglijskij frazeologičeskij slovar'*, 1st edn, Jazyki russkoj kul'tury, Moscow.
- Lyons, J. 1977, *Semantics*, Cambridge University Press, Cambridge.
- MacWhinney, B. 1989, "Competition and lexical categorization" in *Linguistic categorization*, eds. R. Corrigan, F. Eckman & M. Noonan, John Benjamins, Amsterdam, pp. 195-242.
- Makkai, A. 1972, *Idiom structure in English*, Mouton, The Hague.
- Marr, D. & Vaina, L. 1982, "Representation and recognition of the movements of shapes", *Proceedings of the Royal Society of London B*, vol. 214, pp. 501-524.
- Martin, J.R. 2000, "Beyond exchange : APPRAISAL systems in English" in *Evaluation in text*, eds. S. Hunston & G. Thompson, Oxford University Press, Oxford, pp. 142-175.

- Martin, J.R. & Rose, D. 2003, *Working with discourse : meaning beyond the clause*, Continuum, London.
- McCawley, J. 1968, "Concerning the base component of a transformational grammar", *Foundations of Language*, vol. 4, pp. 243-269.
- McElree, B. & Griffith, T. 1995, "Syntactic and thematic processing in sentence comprehension : evidence for a temporal dissociation", *Journal of Experimental Psychology : Learning, Memory, and Cognition*, vol. 21, pp. 134-157.
- McElree, B. & Nordlie, J. 1999, "Literal and figurative interpretations are computed in equal time", *Psychonomic Bulletin & Review*, vol. 6, pp. 486-494.
- McGlone, M.S. 2001, "Concepts as metaphors" in *Understanding figurative language : from metaphors to idioms*, ed. S. Glucksberg, Oxford University Press, Oxford, pp. 90-107.
- McGlone, M.S., Glucksberg, S. & Cacciari, C. 1994, "Semantic productivity and idiom comprehension", *Discourse Processes*, vol. 17, pp. 167-190.
- Mel'čuk, I.A. 2007, "Lexical functions" in *Phraseologie : ein internationales Handbuch zeitgenössischer Forschung = Phraseology : an international handbook of contemporary research*, eds. H. Burger, D. Dobrovol'skij, P. Kühn & N.R. Norick, Mouton De Gruyter, Berlin/New York, pp. 119-131.
- Mel'čuk, I.A. 2004, "Actants in semantics and syntax I : actants in semantics", *Linguistics*, vol. 42, no. 1, pp. 1-66.
- Mel'čuk, I.A. 2004, "Actants in semantics and syntax II : actants in syntax", *Linguistics*, vol. 42, no. 2, pp. 247-291.
- Mel'čuk, I.A. 2001, *Communicative organization in natural language : the semantic-communicative structure of sentences*, J. Benjamins, Amsterdam ; Philadelphia.
- Mel'čuk, I.A. 1998, "Collocations and lexical functions" in *Phraseology : theory, analysis, and applications*, ed. A.P. Cowie, Clarendon Press ; Oxford Univ. Press, Oxford, pp. 23-53.
- Mel'čuk, I.A. 1996, "Lexical functions : a tool for the description of lexical relations in a lexicon" in *Lexical functions in lexicography and natural language processing*, ed. L. Wanner, John Benjamins Publishing Company, Amsterdam, pp. 37-102.
- Mel'čuk, I.A. 1995, "Phrasemes in language and phraseology in linguistics" in *Idioms : structural and psychological perspectives*, eds. M. Everaert, E. van der Linden, A.Y. Schenk & R. Schreuder, Lawrence Erlbaum Associates, Hillsdale New Jersey, pp. 167-232.
- Mel'čuk, I.A. 1981, "Meaning-Text Models", *Annual Review of Anthropology*, vol. 10, pp. 27-62.
- Mel'čuk, I.A. 1974, *Opyt teorii lingvističeskikh modelej "smysl - tekst" : semantika, sintaksis*, Nauka, Moscow.
- Mel'čuk, I.A. 1973, "Linguistic theory and "Meaning-Text" type models" in *Logic, language, and probability; a selection of papers contributed to sections IV, VI, and XI of the Fourth International Congress for Logic, Methodology, and Philosophy of Science, Bucharest, September 1971*, eds. R.J. Bogdan & I. Niiniluoto, D. Reidel Pub. Co, Dordrecht, Holland, Boston, pp. 223-235.



- Mel'čuk, I.A. & Žolkovskij, A.K. 1984, *Tolkovo-kombinatornyj slovar' sovremennogo russkogo âzyka : opyty semantiko-sintaksičeskogo opisaniâ russkoj leksiki*, Wiener Slawistischer Almanach, Vena.
- Mel'čuk, I.A. & Žolkovskij, A.K. 1970, "Towards a functioning Meaning-Text Model of language", *Linguistics*, vol. 57, pp. 10-47.
- Melerovič, A.M. & Mokienko, V.M. 2001, *Frazeologizmy v russkoj reči : slovar'*, 2<sup>nd</sup> edn, Russkie slovari, Moscow.
- Melerovič, G. 1998, "On semantic structure of the proverbs", *Europhras '97 : Phraseology and Paremiology (International Symposium, September 2-5, 1997, Liptovský Ján)*, ed. P. Ďurčo, Akadémia PZ, Bratislava, pp. 225-229.
- Michaelis, L.A. 1998, *Aspectual grammar and past-time reference*, Routledge, London ; New York.
- Michaelis, L.A. & Lambrecht, K. 1996, "The exclamative sentence type in English" in *Conceptual structure, discourse, and language*, ed. A.E. Goldberg, CSLI Publications, Stanford, CA, pp. 375-389.
- Mikone, E. 2000, "Idiomit ja kieliopillistuminen" in *Fennistica fausta in honorem Mauno Koski septuagenarii = Juhlakirja Mauno Kosken 70-vuotispäivän kunniaksi 21.2.2000*, eds. E. Lähdemäki & Y. Bertills, Åbo Akademi, Åbo, pp. 16-32.
- Miličević, J. 2006, "A short guide to the Meaning-Text linguistic theory", *Journal of Koralex*, vol. 8, pp. 187-233.
- Mokienko, V.M. 1980, *Slavânskaâ frazeologijâ*, Moscow.
- Molotov, A.I. 1977, *Osnovy frazeologii russkogo âzyka*, Nauka, Leningrad.
- Moon, R. 1998, *Fixed expressions and idioms in English : a corpus-based approach*, Clarendon Press, Oxford.
- MOT Englanti 4.8 sanakirja, see Kielikone Ltd. 2010.
- Murphy, G.L. 1997, "Reasons to doubt the present evidence for metaphoric representation", *Cognition*, vol. 62, pp. 99-1084.
- Murphy, G.L. 1996, "On metaphoric representation", *Cognition*, vol. 60, pp. 173-204.
- Naciscione, A. 2001, *Phraseological units in discourse : towards applied stylistics*, Latvian Academy of Culture, Riga.
- Nenonen, M. 2007a, "Prototypical idioms : evidence from Finnish", *SKY Journal of Linguistics*, vol. 20, pp. 330-330.
- Nenonen, M. 2007b, "Unique, but not cranberries : idiomatic isolates in Finnish", *Collocations and idioms 1 : papers from the First Nordic Conference on Syntactic Freezes, Joensuu, May 19-20, 2006*, eds. M. Nenonen & S. Niemi, Joensuun yliopisto, Joensuu, pp. 213-226.
- Nenonen, M. 2002, *Idiomit ja leksikko lausekeidiomien syntaktisia, semanttisia ja morfologisia piirteitä suomen kielessä*, Joensuun yliopisto, Joensuu.
- Nenonen, M. 2001a, "Idiomaticity in Finnish", *Nordic and Baltic Morphology. Papers from a NorFA course, Tartu, June 2000*, eds. J. Niemi & J. Heikkinen, University of Joensuu, Joensuu, pp. 57-65.

- Nenonen, M. 2001b, "Suomen kielen idiomaattisista ilmauksista - morfologinen katsaus", *Kielitieteen päivät Oulussa 2000*, eds. H. Sulkala & L. Nissilä, Acta Universitatis Ouluensis, Oulu, pp. 182-190.
- Nenonen, M. & Niemi, J. 2010, "Mismatches between grammatical number and conceptual numerosity : a number decision experiment on collective nouns, number neutralization, pluralia tantum and idiomatic plurals", *Folia Linguistica*, vol. 44, no. 1, pp. 103-126.
- Nenonen, M. & Niemi, J. 2004, "Mismatches between number and numerosity : perceptual salience and Givon's quantity principles", *Papers from the 30th Finnish Conference of Linguistics, Joensuu, May 15-16, 2003*, ed. M. Nenonen, University of Joensuu, Joensuu, pp. 129-134.
- Nenonen, M. & Niemi, J. 1999, "Morphological isolates in idioms : cranberries or real words?", *Brain and Language*, vol. 68, pp. 158-164.
- Nenonen, M., Niemi, J. & Laine, M. 2002, "Representation and processing of idioms : evidence from aphasia", *Journal of Neurolinguistics*, vol. 15, no. 1, pp. 43-58.
- Nenonen, M., Niemi, J. & Laine, M. 2000, "Morphologically isolated words in idioms : a case study with an aphasic", *Language and Verbal Behaviour*, vol. 3, no. 2, pp. 72-79.
- Nenonen, M. & Niemi, S. (eds) 2007, *Collocations and idioms 1 : papers from the First Nordic Conference on Syntactic Freezes, Joensuu, May 19-20, 2006*, Joensuun yliopisto, Joensuu.
- Newmeyer, F.J. 1974, "The regularity of idiom behaviour", *Lingua*, vol. 34, pp. 327-342.
- Nicolas, T. 1995, "Semantics of idiom modification" in *Idioms : structural and psychological perspectives*, ed. M. Everaert, Lawrence Erlbaum Associates, Hillsdale NJ, pp. 233-252.
- Niemi, J., Mulli, J., Nenonen, M., Niemi, S., Nikolaev, A. & Penttilä, E. 2010, "Body-part idioms across languages : lexical analyses of VP body-part idioms in English, German, Swedish, Russian and Finnish" in *Korpora, Web und Datenbanken : Computergestützte Methoden in der modernen Phraseologie und Lexikographie = Corpora, web and databases : computer-based methods in modern phraseology and lexicography*, eds. S. Ptashnyk, E. Hallsteinsdóttir & N. Bubenhofer, Schneider Verlag Hohengehren, Baltmannsweiler, pp. 67-76.
- Niemi, J., Nenonen, M. & Penttilä, E. 1998, "Number as a marker of idiomaticity", *Papers from the 16th Scandinavian Conference of Linguistics, Turku/Åbo, November 14-16, 1996*, ed. T. Haukioja, Åbo Akademi Tryckeri, Åbo, pp. 293-304.
- Nikanne, U. 2008a, "Conceptual Semantics" in *Handbook of pragmatics 12*, eds. J. Östman & J. Verschueren, John Benjamins, Amsterdam.
- Nikanne, U. 2008b, *Micro-modular syntax*, PPT-presentation, 2nd workshop on formal approaches to language as a cognitive system, Turku 10-11 April 2008.
- Nikanne, U. 2006, "Aspectual case marking of object in Finnish", *Research in Language*, vol. 4, pp. 215-242.
- Nikanne, U. 2005a, "Constructions in Conceptual Semantics" in *Construction grammars : cognitive grounding and theoretical extensions*, eds. J. Östman & M. Fried, John Benjamins Publishing, Amsterdam, pp. 191-242.
- Nikanne, U. 2005b, "Path expressions in Finnish and Swedish : the role of Constructions" in *Functional features in language and space : insights from language, perception,*

- categorization, and development*, eds. L. Carlson & E. van der Zee, Oxford University Press, Oxford, pp. 311-327.
- Nikanne, U. 2004, "Ruumiin liikkeet, ruumiillisuusoletus ja urheilu. [Body Movements, embodiment hypothesis and athletics]", *Puhe ja Kieli = Tal och Språk*, vol. 24, pp. 101-110.
- Nikanne, U. 2002, **Kerrokset ja kytkennät : Konseptuaalisen semantiikan perusteet. [Tiers and links : the basics of Conceptual Semantics]** [Internet], Åbo Akademi, Turku. Available from: <<http://web.abo.fi/fak/hf/fin/kurssit/KONSEM/tekstit.htm>> [Accessed 12 Feb 2005]
- Nikanne, U. 2000, "Some restrictions in linguistic expressions of spatial movement" in *Cognitive interfaces : constraints on linking cognitive information*, eds. E. van der Zee & U. Nikanne, Oxford University Press, New York, pp. 77-93.
- Nikanne, U. 1998, "The lexicon and conceptual structure", *Papers from the 16th Scandinavian conference of linguistics*, ed. T. Haukioja, Turun yliopisto, Turku, pp. 305-319.
- Nikanne, U. 1997a, "Lexical conceptual structure and argument linking", *SKY Journal of Linguistics*, pp. 81-118.
- Nikanne, U. 1997b, "Suomen infiniittisten adjunktien temporaalinen tulkinta", *Virittäjä*, no. 3, pp. 338-357.
- Nikanne, U. 1997c, "Locative case adjuncts in Finnish : notes on syntactico-semantic interface", *Nordic Journal of Linguistics*, vol. 20, pp. 155-178.
- Nikanne, U. 1995, "Action tier formation and argument linking", *Studia Linguistica*, vol. 49, pp. 1-31.
- Nikanne, U. 1992, "Metaforien mukana" in *Metafora : ikkuna kieleen, mieleen ja kulttuuriin*, eds. L. Harvilahti, J. Kalliokoski, U. Nikanne & T. Onikki, Suomalaisen kirjallisuuden seura, Helsinki, pp. 60-78.
- Nikanne, U. 1990a, "Possessiivinen kenttä" in *Suomen kielen paikallissijat konseptuaalisessa semantiikassa*, eds. P. Leino et al., Helsingin yliopisto, Helsinki, pp. 183-199.
- Nikanne, U. 1990b, *Zones and tiers : a study of thematic structure*, Suomalaisen kirjallisuuden seura, Helsinki.
- Nikanne, U. & Östman, J. 2006, "Finland-Swedish directionality in Conceptual Semantics and in Construction Grammar : a methodological dialogue" in *A man of measure : Festschrift in honour of Fred Karlsson on his 60th birthday*, eds. M. Suominen et al., SKY Journal of Linguistics, special supplement to vol. 19, SKY, Turku, pp. 66-86.
- Nikanne, U. & van der Zee, E. 2011, in press, "The lexical representation of path-curvature in motion expressions : a three-way path-curvature distinction" in *Motion encoding in language and space*, eds. M. Vulchanova & E. van der Zee, Oxford University Press, Oxford.
- Noh, E. 2000, *Metarepresentation : a relevance-theory approach*, John Benjamins, Amsterdam.
- NS = *Nykysuomen sanakirja*, see Sadeniemi, M. 1970.
- Nunberg, G., Sag, I.A. & Wasow, T. 1994, "Idioms", *Language*, vol. 70, no. 3, pp. 491-538.
- ODEI = *Oxford dictionary of English idioms*, see Cowie, A.P., Mackin, R. & McCaig, I.R. 1993.
- O'Grady, W. 1998, "The syntax of idioms", *Natural Language and Linguistic Theory*, vol. 16, pp. 279-312.

- Ogden, C.K., Richards, I.A., Malinowski, B., Crookshank, F.G. & Postgate, J.P. 1923, *The meaning of meaning : a study of the influence of language upon thought and of the science of symbolism*, K. Paul, Trench, Trubner & co., ltd ; Harcourt, Brace & company, inc, London ; New York.
- Ojanen, J. & Uotila-Arcelli, L. 1977, "Suomen kielen frekventatiivi- ja momentaaniverbien derivaatio", *Virittäjä*, vol. 81, pp. 245-264.
- Omazić, M. 2008, "Processing of idioms and idiom modifications : a view from cognitive linguistics" in *Phraseology : an interdisciplinary perspective*, eds. S. Granger & F. Meunier, John Benjamins Publishing Company, Amsterdam, pp. 67-79.
- Omazić, M. 2007, "Phraseological blends", *Europhras Slovenija 2005 : Frazeologija v jezikoslovju in drugih vedah = Phraseology in linguistics and other branches of science = Phraseologie in der Sprachwissenschaft und anderen Disziplinen = Frazeologija v âzykoznanii i drugih naukah*, eds. E. Kržišnik & W. Eismann, Univerza v Ljubljani, Ljubljana, pp. 99-108.
- Omazić, M. & Delibegović Džanić, N. 2009, "Constraints to mechanisms of idiom modification" in *Phraseologie disziplinär und interdisziplinär*, ed. C. Földes, Gunter Narr Verlag, Tübingen, pp. 211-222.
- Ortony, A. 1988, "Are emotion metaphors conceptual or lexical?", *Cognition and Emotion*, vol. 2, pp. 95-103.
- Ortony, A. (ed) 1979, *Metaphor and thought*, Cambridge University Press, Cambridge.
- Ortony, A., Schallert, D.L., Reynolds, R.E. & Antos, S.J. 1978, "Interpreting metaphors and idioms : some effects of context on comprehension", *Journal of Verbal Learning and Verbal Behaviour*, vol. 17, pp. 465-477.
- Östman, J. 2002, "Sulvan kansan wellerismit konstruktioina" in *Äidinkielen Merkitykset*, eds. I. Herlin, J. Kalliokoski, L. Kotilainen & T. Onikki-Rantajääskö, Suomalaisen Kirjallisuuden Seura, Helsinki, pp. 75-97.
- Östman, J. & Fried, M. (eds) 2005, *Construction grammars : cognitive grounding and theoretical extensions*, John Benjamins, Amsterdam.
- Palacios Martínez, I.M. 1999, "Negative polarity idioms in modern English", *ICAME Journal*, vol. 23, pp. 65-115.
- Parkkinen, J. & Harju, J. 2005, *Aasinsilta ajan hermolla*, Wsoy, Helsinki.
- Partington, A. 1998, *Patterns and meanings : using corpora for English language research and teaching*, John Benjamins, Amsterdam.
- Paulsen, G. 2011, *Causation and dominance : a study of Finnish causative verbs expressing social dominance*, Åbo Akademi university press, Turku.
- Penttilä, E. 2006, *It takes an age to do a Chomsky : idiomaticity and verb phrase constructions in English*, University of Joensuu, Joensuu.
- Penttilä, E., Nenonen, M. & Niemi, J. 1998, "Cultural and biological bases of idioms : a crosslinguistic study" in *Language contact, variation, and change*, eds. J. Niemi, T. Odlin & J. Heikkinen, University of Joensuu, Joensuu, pp. 234-245.
- Permiakov, G.L. 1979, *Ot pogovorki do skazki : zametki po obšej teorii kliše*. [From proverb to folk-tale : notes on the general theory of cliché], Nauka, Moscow.

- Peterson, R.R. & Burgess, C. 1993, "Syntactic and semantic processing during idiom comprehension : neurolinguistic and psycholinguistic dissociations" in *Idioms : processing, structure, and interpretation*, eds. C. Cacciari & P. Tabossi, Erlbaum, Hillsdale, N.J., pp. 201-225.
- Peterson, R.R., Burgess, C., Dell, G.S. & Eberhard, K. 2001, "Dissociation between syntactic and semantic processing during idiom comprehension", *Journal of Experimental Psychology : Learning, Memory, and Cognition*, vol. 27, pp. 1223-1237.
- Petrova, O. 2010, "Idiom Variation in Computer-Mediated Discourse (Google-groups as a source of data for phraseological research in Finnish)" in *Korpora, Web und Datenbanken : Computergestützte Methoden in der modernen Phraseologie und Lexikographie = Corpora, web and databases : computer-based methods in modern phraseology and lexicography*, eds. S. Ptashnyk, E. Hallsteinsdóttir & N. Bubenhofer, Schneider Verlag Hohengehren, Baltmannsweiler, pp. 133-150.
- Petrova, O. 2009, "Negative modality in Finnish phraseological units" in *Phraseologie disziplinar und interdisziplinär*, ed. C. Földes, Gunter Narr Verlag, Tübingen, pp. 237-252.
- Petrova, O. 2007a, "Component hierarchy in phraseological units", *Europhras Slovenija 2005 : Frazelogija v jezikoslovju in drugih vedah = Phraseology in linguistics and other branches of science = Phraseologie in der Sprachwissenschaft und anderen Disziplinen = Frazelogiâ v âzykoznanii i drugih naukah*, eds. E. Kržišnik & W. Eismann, Univerza v Ljubljani, Ljubljana, pp. 109-118.
- Petrova, O. 2007b, "Phraseological component isolation in computer-mediated discourse : a formal approach" in *Collocations and idioms 1 : Papers from the First Nordic Conference on Syntactic Freezes, Joensuu, May 19-20, 2006*, eds. M. Nenonen & S. Niemi, University of Joensuu, Joensuu, pp. 269-281.
- Piaget, J. 1966, *The psychology of intelligence*, Littlefield, Adams & Co., Totowa, N.J.
- Podgurska, B. 1990, *Problemy individual'no-avtorskih transformacij frazeologičeskikh edinic v romanah I. Il'fa i E. Petrova "Dvenadcat' stul'ev" i "Zolotoj telenok" i sposoby ih perevoda na pol'skij jazyk*, Candidate of philological sciences thesis, AN SSSR, In-t rus. jaz., Moscow.
- Pörn, M. 2004, *Suomen tunnekausatiiviverbit ja niiden lausemaisest täydennykset*, Suomalaisen Kirjallisuuden Seura, Helsinki.
- PS = *Suomen kielen perussanakirja*, see Haarala, R. 1990.
- Pulman, S. 1993, "The recognition and interpretation of idioms" in *Idioms : processing, structure, and interpretation*, eds. C. Cacciari & P. Tabossi, Erlbaum, Hillsdale, N.J., pp. 249-270.
- Pustejovsky, J. 2003, "Categories, types, and qualia structure" in *Asymmetry in grammar, volume 1 : syntax and semantics*, ed. A. Di Sciullo, John Benjamins, Amsterdam, pp. 373-393.
- Pustejovsky, J. 2001, "Type construction and the logic of concepts" in *The language of word meaning*, eds. P. Bouillon & F. Busa, Cambridge University Press, Cambridge, pp. 91-123.
- Pustejovsky, J. 1995, *The generative lexicon*, Mit Press, Cambridge MA.

- Rheingold, H. 2000, *The virtual community : homesteading on the electronic frontier*, Rev edn, MIT Press, Cambridge, Mass.
- Ritchie, D. 2004, "Lost in space : metaphors in conceptual integration theory", *Metaphor and Symbol*, vol. 19, pp. 31-50.
- Rosch, E. 1975, "Cognitive representations of semantic categories", *Journal of Experimental Psychology : General*, vol. 104, pp. 192-233.
- Rozental', D.Ė. & Telenkova, M.A. 1985, *Slovar'-spravočnik lingvističeskikh terminov*, 3rd edn, Prosvešenie, Moscow.
- Russian National Corpus 2011, **Russian National Corpus** [Internet]. Available from: <<http://www.ruscorpora.ru>> [Accessed 14 May 2009]
- Sadeniemi, M. 1970, *Nykysuomen sanakirja*, Lyhentämätön kansanp., 3<sup>rd</sup> edn, Wsoy, Porvoo.
- Sadeniemi, M. 1942, "Huomioita kielteisistä ja myönteisistä ilmauksista", *Virittäjä*, vol. 46, pp. 285-289.
- Sananparsikokoelma 2010, **Sananparsikokoelma** [Internet], Kotimaisten kielten tutkimuskeskus, Helsinki. Available from: <[http://kaino.kotus.fi/korpus/sp/meta/sp\\_coll\\_rdf.xml](http://kaino.kotus.fi/korpus/sp/meta/sp_coll_rdf.xml)> [Accessed 25 Jan 2009]
- Šanskij, N.M. 1963, *Frazeologičeskoe sovremennogo russkogo jazyka*, Vysšaja škola, Moscow.
- Schenk, A. 1995, "The syntactic behaviour of idioms" in *Idioms : structural and psychological perspectives*, ed. M. Everaert, Lawrence Erlbaum Associates, Hillsdale NJ, pp. 253-271.
- Schweigert, W.A. & Moates, D.R. 1988, "Familiar idiom comprehension", *Journal of Psycholinguistic Research*, vol. 17, no. 281, pp. 296.
- Searle, J.R. 1979, *Expression and meaning : studies in the theory of speech acts*, Cambridge U.P, Cambridge.
- Sinclair, J. 1991, *Corpus, concordance, collocation*, Oxford Univ. Press, Oxford.
- Sinnemäki, M. 1982, *Lentävien lauseiden sanakirja (Suomen kielen sanakirjat 5)*, Otava, Helsingissä.
- Sköldberg, E. 2004, *Korten på bordet : innehålls- och uttrycksmässig variation hos svenska idiom*, Meijerbergs institut för svensk etymologisk forskning, Göteborg.
- Smith, N.V. 1999, *Chomsky : ideas and ideals*, Cambridge University Press, Cambridge, UK ; New York.
- Solodub, J.P. & Al'brecht, F.B. 2003, *Sovremennyj russkij jazyk : leksika i frazeologija (sopostavitel'nyj aspekt)*, [Modern Russian : lexicon and phraseology (comparative aspect)], Flinta/Nauka, Moscow.
- Sovran, T. 1993, "Metaphor as reconciliation - the logical-semantic basis of metaphorical juxtaposition", *Poetics Today*, vol. 14, no. 1, pp. 25-48.
- Sperber, D. & Wilson, D. 1986, *Relevance : communication and cognition*, Harvard University Press, Cambridge MA.
- Stefanowitsch, A. 2003, "A construction-based approach to indirect speech acts" in *Metonymy and pragmatic inferencing*, eds. K. Panther & L.L. Thornburg, John Benjamins, Amsterdam ; Philadelphia, pp. 105-126.
- Stock, O., Slack, J. & Ortony, A. 1993, "Building castles in the air : some computational and theoretical issues in idiom comprehension" in *Idioms : processing, structure, and*

- interpretation*, eds. C. Cacciari & P. Tabossi, Lawrence Erlbaum and Associates, Hillsdale, N.J., pp. 229-248.
- Stubbs, M. 2001, *Words and phrases : corpus studies of lexical semantics*, Blackwell, Oxford.
- Suihkonen, P. 1994, *Frekventatiividerivaatio suomen murteissa : morfologis-fonologista tarkastelua*, Suomalaisen Kirjallisuuden Seura, Helsinki.
- Summers, D. 1987, *Longman dictionary of contemporary English*, New edn, Longman, Harlow.
- Swinney, D.A. 1979, "Lexical access during sentence comprehension : (re)consideration of context effects", *Journal of Verbal Learning & Verbal Behaviour*, vol. 18, no. 6, pp. 645-659.
- Swinney, D.A. & Cutler, A. 1979, "The access and processing of idiomatic expressions", *Journal of Verbal Learning & Verbal Behaviour*, vol. 18, pp. 523-534.
- Tabossi, P. & Zardon, F. 1993, "The activation of idiomatic meaning in spoken language comprehension" in *Idioms : processing, structure, and interpretation*, eds. C. Cacciari & P. Tabossi, Lawrence Erlbaum and Associates, Hillsdale, N.J., pp. 145-162.
- Takalo, R. & Junttila, T. 1979, **Sananparsia ja sutkauksia Nivalan seudulta** [Internet], KirjastoVirma, Haapajärvi. Available from: <<http://www.kirjastovirma.net/nivala/sutkauksia>> [Accessed 29 Oct 2009]
- Talmy, L. 1985, "Force dynamics in language and thought", *Papers from the 21st Regional Meeting of the Chicago Linguistic Society*, Chicago, pp. 293-337.
- Tanenhaus, M.K., Leiman, J.M. & Seidenberg, M.S. 1979, "Evidence for multiple stages in the processing of ambiguous words in syntactic contexts", *Journal of Verbal Learning & Verbal Behaviour*, vol. 18, pp. 427-441.
- Tanskanen, S. 2006, *Collaborating towards coherence : lexical cohesion in English discourse*, John Benjamins ; Philadelphia, Amsterdam.
- Taylor, J.R. 2003, *Linguistic categorization : prototypes in linguistic theory*, 3rd edn, Oxford University Press, Oxford.
- Teliâ, V.N. 1996, *Ruskaâ frazeologiâ : semantičeskij, pragmaticeskij i lingvokul'turologičeskij aspekty*, Škola "Âziki Russkoj Kul'tury", Moscow.
- Teliâ, V.N. 1990, "Semantika idiom v funkcional'no-parametričeskom otobraženii [Idiom semantics in functional-parametric representation]" in *Frazeografiâ v Mašinnom fonde russkogo âzyka*, ed. V.N. Teliâ, Nauka, Moscow, pp. 32-47.
- Teliâ, V.N. 1986, *Konnotativnyj aspekt semantiki nominativnyh edinic*, Nauka, Moscow.
- Titone, D.A. & Connine, C.M. 1994, "Descriptive norms for 171 idiomatic expressions : familiarity, compositionality, predictability, and literality", *Metaphor & Symbolic Activity*, vol. 9, pp. 247-270.
- Titone, D.A. & Connine, C.M. 1999, "On the compositional and noncompositional nature of idiomatic expressions", *Journal of Pragmatics*, vol. 31, pp. 1655-1674.
- Tronenko, N. 2004, *Regularities in the behaviour of Russian phrasal idioms*, Peter Lang, Oxford.
- Turner, K. 1999, *The semantics/pragmatics interface from different points of view*, Elsevier, Oxford England ; New York.
- Vainio, S. & Nenonen, M. 2007, "Reading idioms in Finnish : an eye-tracking study" in *Collocations and idioms 1 : Papers from the First Nordic Conference on Syntactic Freezes*,

- Joensuu, May 19-20, 2006, eds. M. Nenonen & S. Niemi, University of Joensuu, Joensuu, pp. 363-377.
- van de Voort, M. & Vonk, W. 1995, "You don't die immediately when you kick an empty bucket : a processing view on semantic and syntactic characters of idioms" in *Idioms : structural and psychological perspectives*, eds. M. Everaert, E. van der Linden, A. Schenk & R. Schreuder, Lawrence Erlbaum Associates, Hillsdale, New Jersey, pp. 283-300.
- van der Auwera, J. 1999, "On the semantic and pragmatic polyfunctionality of modal verbs" in *The semantics/pragmatics interface from different points of view*, ed. K. Turner, Elsevier, Oxford England ; New York, pp. 49-64.
- van der Linden, E. 1993, *A categorial, computational theory of idioms*, LEd, Utrecht.
- van der Zee, E. 2000, "Why we can talk about bulging barrels and spinning spirals : curvature representation in the lexical interface" in *Cognitive interfaces : constraints on linking cognitive information*, eds. E. van der Zee & U. Nikanne, Oxford University Press, New York, pp. 143-182.
- van der Zee, E. 1996, *Spatial knowledge and spatial language : a theoretical and empirical investigation*, ISOR, Utrecht.
- van der Zee, E. & Nikanne, U. 2000, *Cognitive interfaces : constraints on linking cognitive information*, Oxford University Press, New York.
- van der Zee, E. & Nikanne, U. 2000, "Introducing cognitive interfaces and constraints on linking cognitive information" in *Cognitive interfaces : constraints on linking cognitive information*, eds. E. van der Zee & U. Nikanne, Oxford University Press, New York, pp. 1-17.
- van der Zee, E., Nikanne, U. & Sassenberg, U. 2010, "Grain levels in English path curvature descriptions and accompanying iconic gestures", *Journal of Spatial Information Science*, vol. 1, pp. 95-113.
- van der Zee, E. & Slack, J.M. 2003, *Representing direction in language and space*, Oxford University Press, Oxford.
- van Gestel, F. 1995, "En bloc insertion" in *Idioms : structural and psychological perspectives*, eds. M. Everaert, E. van der Linden, A. Schenk & R. Schreuder, Lawrence Erlbaum Associates, Hillsdale, New Jersey, pp. 75-96.
- Vega Moreno, R.E. 2007, *Creativity and convention : the pragmatics of everyday figurative speech*, John Benjamins, Amsterdam ; Philadelphia.
- Vilkuna, M. 1996, *Suomen lauseopin perusteet*, Edita, Helsinki.
- Vinogradov, V.V. 1977, "Osnovnyje ponjatija russkoj frazeologii kak lingvističeskoj discipliny" in V. V. Vinogradov. *Izbrannye trudy. Leksikologija i leksokografija*, ed. V.G. Kostomarov, Nauka, Moscow, pp. 118-139.
- Volosevič, S.P. 1989, *Zamena komponentov frazeologičeskih edinic v aspekte nominacii i komunikacii (na materiale sovremennogo anglijskogo âzyka*, Candidate of philological sciences thesis, Maurice Thorez Moscow State Pedagogical Institute of Foreign Languages, Moscow.
- Vološinov, V.N. 1973, *Marxism and the philosophy of language*. Translated by L. Matejka & I.R. Titunik. Seminar Press, New York.



- Wanner, L. (ed) 2007, *Selected lexical and grammatical issues in the meaning-text theory : in honour of Igor Mel'čuk*, John Benjamins, Amsterdam ; Philadelphia.
- Wanner, L. (ed) 1996, *Lexical functions in lexicography and natural language processing*, John Benjamins, Amsterdam ; Philadelphia.
- Wasow, T., Sag, I.A. & Nunberg, G. 1983, "Idioms : an interim report", *Proceedings of the XIIIth International Congress of Linguists, August 29 -September 4, 1982, Tokyo* CIPL, Tokyo, pp. 102-115.
- Weinreich, U. 1969, "Problems in the analysis of idioms" in *Substance and structure of language : lectures delivered before the Linguistic institute of the Linguistics society of America, University of California, Los Angeles, June 17 - August 12, 1966*, ed. J. Puhvel, University of California Press, Berkeley CA, pp. 23-81.
- White, P.R.R. 2005, **The Appraisal website : homepage** [Internet], Verbosity Enterprises, Birmingham. Available from: <[http://www.grammatics.com/appraisal /index.html](http://www.grammatics.com/appraisal/index.html)> [Accessed 10 April 2008]
- White, P.R.R. 2002, "Appraisal - the language of evaluation and stance" in *Handbook of Pragmatics*, eds. J. Verschueren, J. Östman, J. Blommaert & C. Bulcaen, John Benjamins, Amsterdam, pp. 1-27.
- Wierzbicka, A. 1967, "On the semantics of verbal aspect in Polish" in *To honor Roman Jakobson : essays on the occasion of his seventieth birthday 11 october 1966*, vol. 3 Mouton, The Hague, pp. 2231-2249.
- Wiik, K. 1975, "Suomen frekventatiivi ja kontinuatiiivi", *Virittäjä*, vol. 79, pp. 153-167.
- Wittgenstein, L. 1953, *Philosophical investigations. Translated by G. E. M. Anscombe. (Philosophische Untersuchungen.) Eng. & Ger*, Oxford, Basil Blackwell.
- Wray, A. 2002, *Formulaic language and the lexicon*, Cambridge University Press, Cambridge.
- Zhang, Q. 1998, "Fuzziness – vagueness – generality – ambiguity", *Journal of Pragmatics*, vol. 29, no. 1, pp. 13-31.
- Žukov, A.V. & Žukov, K.A. 2003, "Semantika neopredeljennosti : o slovah i frazeologizmah s razmytyim i širokim značeniem", *Vestnik Novgorodskogo gosudarstvennogo universiteta*, vol. 25, pp. 106-110.
- Žukov, A.V. 1999, "Zametki o frazeologičeskom značenii", *Uchenye Zapiski Novgorodskogo Gosudarstvennogo Universiteta [Novgorod State University Scientific Papers]* [Internet]. Available from: <<http://www.admin.novsu.ac.ru/uni/scpapers.nsf/publications/D949A515360B4BCBC32567BA003568C1>> [Accessed 10 Sept 2005]
- Žukov, V.P. 1978, *Semantika frazeologičeskikh oborotov*, Prosveščenie, Moscow.
- Žukov, V.P. 1986, *Russkaâ frazeologiâ : učebnoe posobie*, Vyššaâ škola, Moscow.
- Zwicky, A.M. 1973, "Review article of Jan G. Kooij. *Ambiguity in natural language : an investigation of certain problems in its linguistic description*", *Lingua*, vol. 32, pp. 95-118.
- Zwicky, A.M. & Sadock, J.M. 1975, "Ambiguity tests and how to fail them" in *Syntax and Semantics 4*, ed. J.P. Kimball, Academic Press, New York, pp. 1-36.

## Abbreviations and symbols

[ ]	the boundaries of a conceptual constituent, e.g. a thematic argument, e.g. [HOUSE]
ABL	ablatiivi, e.g. <i>lattia-lta</i> 'from the floor'
ADE	adessive case, e.g. <i>lattia-lla</i> 'on the floor'
ALL	allative case, e.g. <i>lattia-lle</i> 'to the floor'
ARG	thematic argument
CS	(rule-based) conceptual structure
CS/PU	conceptual structure licensed by PU
Dfl	(rule-based) default linking
Dfl(PU/C)	default linking licensed by PU/construction
DVtrN	deverbal noun, e.g. <i>heittäminen</i> 'throwing'
ELA	elative case, e.g. <i>talo-sta</i> 'from a house'
#ENTITY#	projected-world entity
#ENTITY#/PU	projected-world entity licensed by PU
Fxd	(rule-based) fixed linking
Fxd(PU/C)	fixed linking licensed by PU/construction
GEN	genitive case, e.g. <i>talo-n</i> 'of a house'
GG	Google Groups
HHS	<i>HEITTÄÄ HELMIÄ SIOILLE</i> 'throw pearls to pigs'
ILL	illative case, e.g. <i>talo-on</i> 'into a house'
INE	inessive case, e.g. <i>talo-ssa</i> 'in a house'
IPC	isolated phraseological constituent
KSH	<i>KUIN SIKA HOPEALUSIKASTA</i> 'like a pig about a silver spoon'
MS	morphological structure
NDfl	non-default linking
NDfl(PU/C)	non-default linking licensed by PU/construction
PCI	phraseological constituent isolation
PS	phonological structure
PTV	partitive case, e.g. <i>talo-a</i> 'house'
PU	phraseological unit
REF	referential index (tier)
SEPAR	separative local case, e.g. elative
#SITUATION#	projected-world situation
#SITUATION#/PU	projected-world situation licensed by PU
SS	syntactic structure
TERM	terminative local case, e.g. illative
Vintr	intransitive verb, e.g. <i>mennä</i> 'go'
Vtr	transitive verb, e.g. <i>heittää</i> 'throw'
X <sub>ALL PL</sub>	lexical item X in allative case plural form
<X>	'X is optional'
{X/Y}	'X and Y are alternatives'
X <sup>α</sup>	'X binds α'
X -/→ Y	'Y is independent of X; Y is not selected by X directly or indirectly'
X <sub>1</sub> , Y <sub>1</sub>	X corresponds to Y

## Structure and variation of phraseological units

The present study deals with phraseological units (PUs), i.e. conventionalized relatively stable multiword items of a given language exhibiting various kinds of irregularities in their structure. It demonstrates what difficulties their structure and variation raises for linguistic theory and offers an alternative way, in which these items can be approached within the framework of Conceptual Semantics and the Tiernet model. The book addresses i.a. the problems of literal vs. idiomatic meaning, analyzability and base form of PUs. PU variants are described in terms of a connectionist network model called construction family. A systematic formalized description of PUs is developed on the basis of a usage-based study of the Finnish PU *HEITTÄÄ HELMIÄ SIOILLE* 'throw pearls to pigs'.

Åbo Akademi University Press  
ISBN 978-951-765-582-8



9 789517 655828

